```
Connecting via Winsock to STN


Welcome to STN International!  Enter x:x

LOGINID:ssspta1653hxp

PASSWORD:
TERMINAL (ENTER 1, 2, 3, OR ?):2


* * * * * * * * * *   Welcome to STN International   * * * * * * * * * *

 NEWS  1           Web Page URLs for STN Seminar Schedule - N. America
 NEWS  2           "Ask CAS" for self-help around the clock
 NEWS  3  JAN 27   Source of Registration (SR) information in REGISTRY updated
                   and searchable
 NEWS  4  JAN 27   A new search aid, the Company Name Thesaurus, available in
                   CA/CAplus
 NEWS  5  FEB 05   German (DE) application and patent publication number format
                   changes
 NEWS  6  MAR 03   MEDLINE and LMEDLINE reloaded
 NEWS  7  MAR 03   MEDLINE file segment of TOXCENTER reloaded
 NEWS  8  MAR 03   FRANCEPAT now available on STN
 NEWS  9  MAR 29   Pharmaceutical Substances (PS) now available on STN
 NEWS 10  MAR 29   WPIFV now available on STN
 NEWS 11  MAR 29   No connect hour charges in WPIFV until May 1, 2004
 NEWS 12  MAR 29   New monthly current-awareness alert (SDI) frequency in RAPRA
 NEWS 13  APR 26   PROMT: New display field available
 NEWS 14  APR 26   IFIPAT/IFIUDB/IFICDB: New super search and display field
                   available
 NEWS 15  APR 26   LITALERT now available on STN
 NEWS 16  APR 27   NLDB: New search and display fields available

 NEWS EXPRESS  MARCH 31 CURRENT WINDOWS VERSION IS V7.00A, CURRENT
               MACINTOSH VERSION IS V6.0c(ENG) AND V6.0Jc(JP),
               AND CURRENT DISCOVER FILE IS DATED 26 APRIL 2004
 NEWS HOURS    STN Operating Hours Plus Help Desk Availability
 NEWS INTER    General Internet Information
 NEWS LOGIN    Welcome Banner and News Items
 NEWS PHONE    Direct Dial and Telecommunication Network Access to STN
 NEWS WWW      CAS World Wide Web Site (general information)

Enter NEWS followed by the item number or name to see news on that
specific topic.

  All use of STN is subject to the provisions of the STN Customer
  agreement.  Please note that this agreement limits use to scientific
  research.  Use for software development or design or implementation
  of commercial gateways or other similar uses is prohibited and may
  result in loss of user privileges and other penalties.

  * * * * * * * * * * * * * * STN  Columbus  * * * * * * * * * * * * * * *

FILE 'HOME' ENTERED AT 14:17:20 ON 29 APR 2004

=> file medline, uspatful, dgene, embase, wpids, jicst
COST IN U.S. DOLLARS                                SINCE FILE      TOTAL
                                                       ENTRY    SESSION
FULL ESTIMATED COST                                     0.63       0.63
```

FILE 'MEDLINE' ENTERED AT 14:18:46 ON 29 APR 2004

FILE 'USPATFULL' ENTERED AT 14:18:46 ON 29 APR 2004
CA INDEXING COPYRIGHT (C) 2004 AMERICAN CHEMICAL SOCIETY (ACS)

FILE 'DGENE' ENTERED AT 14:18:46 ON 29 APR 2004
COPYRIGHT (C) 2004 THOMSON DERWENT

FILE 'EMBASE' ENTERED AT 14:18:46 ON 29 APR 2004
COPYRIGHT (C) 2004 Elsevier Inc. All rights reserved.

FILE 'WPIDS' ENTERED AT 14:18:46 ON 29 APR 2004
COPYRIGHT (C) 2004 THOMSON DERWENT

FILE 'JICST-EPLUS' ENTERED AT 14:18:46 ON 29 APR 2004
COPYRIGHT (C) 2004 Japan Science and Technology Agency (JST)

=> s myelogenous leukemia
L1 22800 MYELOGENOUS LEUKEMIA

=> s myeloid cell proliferation
L2 142 MYELOID CELL PROLIFERATION

=> s "HOIPS I"
L3 18 "HOIPS I"

=> s l3 and l2
L4 0 L3 AND L2

=> s l3 and l1
L5 2 L3 AND L1

=> d l5 ti abs ibib tot

L5 ANSWER 1 OF 2 USPATFULL on STN
TI Human oncogene induced secreted protein I
AB The present invention relates to a novel protein, the Human Oncogene Induced Secreted Protein I ("**HOIPS I**") protein. In particular, isolated nucleic acid molecules are provided encoding the human **HOIPS I** protein. **HOIPS I** polypeptides are also provided as are vectors, host cells and recombinant methods for producing the same. Also provided are diagnostic methods for detecting abnormal cell proliferation and differentiation disorders and therapeutic methods for treating the same.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

ACCESSION NUMBER:	2002:221411 USPATFULL		
TITLE:	Human oncogene induced secreted protein I		
INVENTOR(S):	Olsen, Henrik S., Gaithersburg, MD, UNITED STATES Ruben, Steven M., Olney, MD, UNITED STATES		
PATENT ASSIGNEE(S):	Human Genome Sciences, Inc. (U.S. corporation)		
	NUMBER	KIND	DATE
PATENT INFORMATION:	US 2002119552	A1	20020829
APPLICATION INFO.:	US 2001-899917	A1	20010709 (9)
RELATED APPLN. INFO.:	Division of Ser. No. US 1997-994962, filed on 19 Dec 1997, PATENTED		
	NUMBER		DATE
PRIORITY INFORMATION:	US 1996-33869P		19961220 (60)
	US 1997-37388P		19970207 (60)
DOCUMENT TYPE:	Utility		

FILE SEGMENT: APPLICATION
LEGAL REPRESENTATIVE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C., 1100 NEW YORK AVENUE, N.W., SUITE 600, WASHINGTON, DC, 20005-3934
NUMBER OF CLAIMS: 16
EXEMPLARY CLAIM: 1
NUMBER OF DRAWINGS: 4 Drawing Page(s)
LINE COUNT: 2059
CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L5 ANSWER 2 OF 2 USPATFULL on STN
TI Human oncogene induced secreted protein I
AB The present invention relates to a novel protein, the Human Oncogene Induced Secreted Protein I ("**HOIPS I**") protein. In particular, isolated nucleic acid molecules are provided encoding the human **HOIPS I** protein. **HOIPS I** polypeptides are also provided as are vectors, host cells and recombinant methods for producing the same. Also provided are diagnostic methods for detecting abnormal cell proliferation and differentiation disorders and therapeutic methods for treating the same.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.
ACCESSION NUMBER: 2001:147697 USPATFULL
TITLE: Human oncogene induced secreted protein I
INVENTOR(S): Olsen, Henrik S., Gaithersburg, MD, United States
Ruben, Steven M., Olney, MD, United States
PATENT ASSIGNEE(S): Human Genome Sciences, Inc., Rockville, MD, United States (U.S. corporation)

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 6284486	B1	20010904
APPLICATION INFO.:	US 1997-994962		19971219 (8)

DOCUMENT TYPE: Utility
FILE SEGMENT: GRANTED
PRIMARY EXAMINER: Carlson, Karen Cochrane
LEGAL REPRESENTATIVE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
NUMBER OF CLAIMS: 69
EXEMPLARY CLAIM: 1
NUMBER OF DRAWINGS: 4 Drawing Figure(s); 4 Drawing Page(s)
LINE COUNT: 1994
CAS INDEXING IS AVAILABLE FOR THIS PATENT.

=> d his

(FILE 'HOME' ENTERED AT 14:17:20 ON 29 APR 2004)

FILE 'MEDLINE, USPATFULL, DGENE, EMBASE, WPIDS, JICST-EPLUS' ENTERED AT 14:18:46 ON 29 APR 2004
L1 22800 S MYELOGENOUS LEUKEMIA
L2 142 S MYELOID CELL PROLIFERATION
L3 18 S "HOIPS I"
L4 0 S L3 AND L2
L5 2 S L3 AND L1

=> d l3 ti abs ibib tot

L3 ANSWER 1 OF 18 USPATFULL on STN
TI Human oncogene induced secreted protein I
AB The present invention relates to a novel protein, the Human Oncogene Induced Secreted Protein I ("**HOIPS I**") protein. In particular, isolated nucleic acid molecules are provided encoding the human **HOIPS I** protein. **HOIPS I**

polypeptides are also provided as are vectors, host cells and recombinant methods for producing the same. Also provided are diagnostic methods for detecting abnormal cell proliferation and differentiation disorders and therapeutic methods for treating the same.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

ACCESSION NUMBER:	2002:221411 USPATFULL
TITLE:	Human oncogene induced secreted protein I
INVENTOR(S):	Olsen, Henrik S., Gaithersburg, MD, UNITED STATES Ruben, Steven M., Olney, MD, UNITED STATES
PATENT ASSIGNEE(S):	Human Genome Sciences, Inc. (U.S. corporation)

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 2002119552	A1	20020829
APPLICATION INFO.:	US 2001-899917	A1	20010709 (9)
RELATED APPLN. INFO.:	Division of Ser. No. US 1997-994962, filed on 19 Dec 1997, PATENTED		

	NUMBER	DATE
PRIORITY INFORMATION:	US 1996-33869P	19961220 (60)
	US 1997-37388P	19970207 (60)

DOCUMENT TYPE:	Utility
FILE SEGMENT:	APPLICATION
LEGAL REPRESENTATIVE:	STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C., 1100 NEW YORK AVENUE, N.W., SUITE 600, WASHINGTON, DC, 20005-3934
NUMBER OF CLAIMS:	16
EXEMPLARY CLAIM:	1
NUMBER OF DRAWINGS:	4 Drawing Page(s)
LINE COUNT:	2059

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L3 ANSWER 2 OF 18 USPATFULL on STN

TI Human oncogene induced secreted protein I

AB The present invention relates to a novel protein, the Human Oncogene Induced Secreted Protein I ("**HOIPS I**") protein. In particular, isolated nucleic acid molecules are provided encoding the human **HOIPS I** protein. **HOIPS I** polypeptides are also provided as are vectors, host cells and recombinant methods for producing the same. Also provided are diagnostic methods for detecting abnormal cell proliferation and differentiation disorders and therapeutic methods for treating the same.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

ACCESSION NUMBER:	2001:147697 USPATFULL
TITLE:	Human oncogene induced secreted protein I
INVENTOR(S):	Olsen, Henrik S., Gaithersburg, MD, United States Ruben, Steven M., Olney, MD, United States
PATENT ASSIGNEE(S):	Human Genome Sciences, Inc., Rockville, MD, United States (U.S. corporation)

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 6284486	B1	20010904
APPLICATION INFO.:	US 1997-994962		19971219 (8)

DOCUMENT TYPE:	Utility
FILE SEGMENT:	GRANTED
PRIMARY EXAMINER:	Carlson, Karen Cochrane
LEGAL REPRESENTATIVE:	Sterne, Kessler, Goldstein & Fox P.L.L.C.
NUMBER OF CLAIMS:	69
EXEMPLARY CLAIM:	1
NUMBER OF DRAWINGS:	4 Drawing Figure(s); 4 Drawing Page(s)

LINE COUNT: 1994
CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L3 ANSWER 3 OF 18 DGENE COPYRIGHT 2004 THOMSON DERWENT on STN
TI New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia
AN AAW69220 Protein DGENE
AB This sequence is the human oncogene induced secreted protein I (**HOIPS I**) of the invention. **HOIPS I** can be used for treating cell proliferative diseases, particularly cancers, such as acute and chronic myelogenous leukaemias. The products can also be used for detection and diagnosis of a cell proliferative or cell differentiation disorders.
ACCESSION NUMBER: AAW69220 Protein DGENE
TITLE: New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia
INVENTOR: Olsen H S; Ruben S M
PATENT ASSIGNEE: (HUMA-N)HUMAN GENOME SCI INC.
PATENT INFO: WO 9828421 A1 19980702 71p
APPLICATION INFO: WO 1997-US23547 19971219
PRIORITY INFO: US 1997-37388 19970207
US 1996-33869 19961220
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 1998-377652 [32]
CROSS REFERENCES: N-PSDB: AAV44745
DESCRIPTION: Human oncogene induced secreted protein I.

L3 ANSWER 4 OF 18 DGENE COPYRIGHT 2004 THOMSON DERWENT on STN
TI New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia
AN AAV44751 cDNA DGENE
AB This sequence represents an expressed sequence tag, which is specifically stated as not being contained within the DNA of the invention. The DNA of the invention encodes the human oncogene induced secreted protein I (**HOIPS I**). **HOIPS I** can be used for treating cell proliferative diseases, particularly cancers, such as acute and chronic myelogenous leukaemias. The products can also be used for detection and diagnosis of a cell proliferative or cell differentiation disorders.
ACCESSION NUMBER: AAV44751 cDNA DGENE
TITLE: New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia
INVENTOR: Olsen H S; Ruben S M
PATENT ASSIGNEE: (HUMA-N)HUMAN GENOME SCI INC.
PATENT INFO: WO 9828421 A1 19980702 71p
APPLICATION INFO: WO 1997-US23547 19971219
PRIORITY INFO: US 1997-37388 19970207
US 1996-33869 19961220
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 1998-377652 [32]
DESCRIPTION: Expressed sequence tag C02431.

L3 ANSWER 5 OF 18 DGENE COPYRIGHT 2004 THOMSON DERWENT on STN
TI New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia

AN AAV44750 cDNA DGENE
AB This sequence represents an expressed sequence tag, which is specifically stated as not being contained within the DNA of the invention. The DNA of the invention encodes the human oncogene induced secreted protein I (**HOIPS I**). **HOIPS I** can be used for
treating cell proliferative diseases, particularly cancers, such as acute and chronic myelogenous leukaemias. The products can also be used for detection and diagnosis of a cell proliferative or cell differentiation disorders.
ACCESSION NUMBER: AAV44750 cDNA DGENE
TITLE: New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia
INVENTOR: Olsen H S; Ruben S M
PATENT ASSIGNEE: (HUMA-N)HUMAN GENOME SCI INC.
PATENT INFO: WO 9828421 A1 19980702 71p
APPLICATION INFO: WO 1997-US23547 19971219
PRIORITY INFO: US 1997-37388 19970207
US 1996-33869 19961220
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 1998-377652 [32]
DESCRIPTION: Expressed sequence tag T84854.

L3 ANSWER 6 OF 18 DGENE COPYRIGHT 2004 THOMSON DERWENT on STN
TI New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia
AN AAV44749 cDNA DGENE
AB This sequence represents an expressed sequence tag, which is specifically stated as not being contained within the DNA of the invention. The DNA of the invention encodes the human oncogene induced secreted protein I (**HOIPS I**). **HOIPS I** can be used for
treating cell proliferative diseases, particularly cancers, such as acute and chronic myelogenous leukaemias. The products can also be used for detection and diagnosis of a cell proliferative or cell differentiation disorders.
ACCESSION NUMBER: AAV44749 cDNA DGENE
TITLE: New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia
INVENTOR: Olsen H S; Ruben S M
PATENT ASSIGNEE: (HUMA-N)HUMAN GENOME SCI INC.
PATENT INFO: WO 9828421 A1 19980702 71p
APPLICATION INFO: WO 1997-US23547 19971219
PRIORITY INFO: US 1997-37388 19970207
US 1996-33869 19961220
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 1998-377652 [32]
DESCRIPTION: Expressed sequence tag T92475.

L3 ANSWER 7 OF 18 DGENE COPYRIGHT 2004 THOMSON DERWENT on STN
TI New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia
AN AAV44748 cDNA DGENE
AB This sequence represents an expressed sequence tag, which is specifically stated as not being contained within the DNA of the invention. The DNA of the invention encodes the human oncogene induced secreted protein I (**HOIPS I**). **HOIPS I** can be used for
treating cell proliferative diseases, particularly cancers, such as acute

and chronic myelogenous leukaemias. The products can also be used for detection and diagnosis of a cell proliferative or cell differentiation disorders.

ACCESSION NUMBER: AAV44748 cDNA DGENE
TITLE: New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia
INVENTOR: Olsen H S; Ruben S M
PATENT ASSIGNEE: (HUMA-N)HUMAN GENOME SCI INC.
PATENT INFO: WO 9828421 A1 19980702 71p
APPLICATION INFO: WO 1997-US23547 19971219
PRIORITY INFO: US 1997-37388 19970207
US 1996-33869 19961220
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 1998-377652 [32]
DESCRIPTION: Expressed sequence tag T91708.

L3 ANSWER 8 OF 18 DGENE COPYRIGHT 2004 THOMSON DERWENT on STN
TI New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia
AN AAV44758 DNA DGENE
AB This sequence is a PCR primer for DNA encoding the human oncogene induced secreted protein I (**HOIPS I**) of the invention.
HOIPS I can be used for treating cell proliferative diseases, particularly cancers, such as acute and chronic myelogenous leukaemias. The products can also be used for detection and diagnosis of a cell proliferative or cell differentiation disorders.

ACCESSION NUMBER: AAV44758 DNA DGENE
TITLE: New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia
INVENTOR: Olsen H S; Ruben S M
PATENT ASSIGNEE: (HUMA-N)HUMAN GENOME SCI INC.
PATENT INFO: WO 9828421 A1 19980702 71p
APPLICATION INFO: WO 1997-US23547 19971219
PRIORITY INFO: US 1997-37388 19970207
US 1996-33869 19961220
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 1998-377652 [32]
DESCRIPTION: Primer for Human oncogene induced secreted protein I.

L3 ANSWER 9 OF 18 DGENE COPYRIGHT 2004 THOMSON DERWENT on STN
TI New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia
AN AAV44757 DNA DGENE
AB This sequence is a PCR primer for DNA encoding the human oncogene induced secreted protein I (**HOIPS I**) of the invention.
HOIPS I can be used for treating cell proliferative diseases, particularly cancers, such as acute and chronic myelogenous leukaemias. The products can also be used for detection and diagnosis of a cell proliferative or cell differentiation disorders.

ACCESSION NUMBER: AAV44757 DNA DGENE
TITLE: New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia
INVENTOR: Olsen H S; Ruben S M
PATENT ASSIGNEE: (HUMA-N)HUMAN GENOME SCI INC.

PATENT INFO: WO 9828421 A1 19980702 71p
APPLICATION INFO: WO 1997-US23547 19971219
PRIORITY INFO: US 1997-37388 19970207
US 1996-33869 19961220
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 1998-377652 [32]
DESCRIPTION: Primer for Human oncogene induced secreted protein I.

L3 ANSWER 10 OF 18 DGENE COPYRIGHT 2004 THOMSON DERWENT on STN
TI New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia
AN AAV44756 DNA DGENE
AB This sequence is a PCR primer for DNA encoding the human oncogene induced secreted protein I (**HOIPS I**) of the invention.
HOIPS I can be used for treating cell proliferative diseases, particularly cancers, such as acute and chronic myelogenous leukaemias. The products can also be used for detection and diagnosis of a cell proliferative or cell differentiation disorders.
ACCESSION NUMBER: AAV44756 DNA DGENE
TITLE: New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia
INVENTOR: Olsen H S; Ruben S M
PATENT ASSIGNEE: (HUMA-N)HUMAN GENOME SCI INC.
PATENT INFO: WO 9828421 A1 19980702 71p
APPLICATION INFO: WO 1997-US23547 19971219
PRIORITY INFO: US 1997-37388 19970207
US 1996-33869 19961220
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 1998-377652 [32]
DESCRIPTION: Primer for Human oncogene induced secreted protein I.

L3 ANSWER 11 OF 18 DGENE COPYRIGHT 2004 THOMSON DERWENT on STN
TI New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia
AN AAV44755 DNA DGENE
AB This sequence is a PCR primer for DNA encoding the human oncogene induced secreted protein I (**HOIPS I**) of the invention.
HOIPS I can be used for treating cell proliferative diseases, particularly cancers, such as acute and chronic myelogenous leukaemias. The products can also be used for detection and diagnosis of a cell proliferative or cell differentiation disorders.
ACCESSION NUMBER: AAV44755 DNA DGENE
TITLE: New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia
INVENTOR: Olsen H S; Ruben S M
PATENT ASSIGNEE: (HUMA-N)HUMAN GENOME SCI INC.
PATENT INFO: WO 9828421 A1 19980702 71p
APPLICATION INFO: WO 1997-US23547 19971219
PRIORITY INFO: US 1997-37388 19970207
US 1996-33869 19961220
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 1998-377652 [32]
DESCRIPTION: Primer for Human oncogene induced secreted protein I.

L3 ANSWER 12 OF 18 DGENE COPYRIGHT 2004 THOMSON DERWENT on STN

TI New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia
AN AAV44754 DNA DGENE
AB This sequence is a PCR primer for DNA encoding the human oncogene induced secreted protein I (**HOIPS I**) of the invention.
HOIPS I can be used for treating cell proliferative diseases, particularly cancers, such as acute and chronic myelogenous leukaemias. The products can also be used for detection and diagnosis of a cell proliferative or cell differentiation disorders.
ACCESSION NUMBER: AAV44754 DNA DGENE
TITLE: New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia
INVENTOR: Olsen H S; Ruben S M
PATENT ASSIGNEE: (HUMA-N) HUMAN GENOME SCI INC.
PATENT INFO: WO 9828421 A1 19980702 71p
APPLICATION INFO: WO 1997-US23547 19971219
PRIORITY INFO: US 1997-37388 19970207
US 1996-33869 19961220
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 1998-377652 [32]
DESCRIPTION: Primer for Human oncogene induced secreted protein I.

L3 ANSWER 13 OF 18 DGENE COPYRIGHT 2004 THOMSON DERWENT on STN
TI New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia
AN AAV44753 DNA DGENE
AB This sequence is a PCR primer for DNA encoding the human oncogene induced secreted protein I (**HOIPS I**) of the invention.
HOIPS I can be used for treating cell proliferative diseases, particularly cancers, such as acute and chronic myelogenous leukaemias. The products can also be used for detection and diagnosis of a cell proliferative or cell differentiation disorders.
ACCESSION NUMBER: AAV44753 DNA DGENE
TITLE: New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia
INVENTOR: Olsen H S; Ruben S M
PATENT ASSIGNEE: (HUMA-N) HUMAN GENOME SCI INC.
PATENT INFO: WO 9828421 A1 19980702 71p
APPLICATION INFO: WO 1997-US23547 19971219
PRIORITY INFO: US 1997-37388 19970207
US 1996-33869 19961220
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 1998-377652 [32]
DESCRIPTION: Primer for Human oncogene induced secreted protein I.

L3 ANSWER 14 OF 18 DGENE COPYRIGHT 2004 THOMSON DERWENT on STN
TI New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia
AN AAV44752 DNA DGENE
AB This sequence is a PCR primer for DNA encoding the human oncogene induced secreted protein I (**HOIPS I**) of the invention.
HOIPS I can be used for treating cell proliferative diseases, particularly cancers, such as acute and chronic myelogenous leukaemias. The products can also be used for detection and diagnosis of a cell proliferative or cell differentiation disorders.

ACCESSION NUMBER: AAV44752 DNA DGENE
TITLE: New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia
INVENTOR: Olsen H S; Ruben S M
PATENT ASSIGNEE: (HUMA-N) HUMAN GENOME SCI INC.
PATENT INFO: WO 9828421 A1 19980702 71p
APPLICATION INFO: WO 1997-US23547 19971219
PRIORITY INFO: US 1997-37388 19970207
US 1996-33869 19961220
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 1998-377652 [32]
DESCRIPTION: Primer for Human oncogene induced secreted protein I.

L3 ANSWER 15 OF 18 DGENE COPYRIGHT 2004 THOMSON DERWENT on STN
TI New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia
AN AAV44746 cDNA DGENE
AB This sequence represents an expressed sequence tag, which is specifically stated as not being contained within the DNA of the invention. The DNA of the invention encodes the human oncogene induced secreted protein I (**HOIPS I**). **HOIPS I** can be used for treating cell proliferative diseases, particularly cancers, such as acute and chronic myelogenous leukaemias. The products can also be used for detection and diagnosis of a cell proliferative or cell differentiation disorders.
ACCESSION NUMBER: AAV44746 cDNA DGENE
TITLE: New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia
INVENTOR: Olsen H S; Ruben S M
PATENT ASSIGNEE: (HUMA-N) HUMAN GENOME SCI INC.
PATENT INFO: WO 9828421 A1 19980702 71p
APPLICATION INFO: WO 1997-US23547 19971219
PRIORITY INFO: US 1997-37388 19970207
US 1996-33869 19961220
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 1998-377652 [32]
DESCRIPTION: Expressed sequence tag.

L3 ANSWER 16 OF 18 DGENE COPYRIGHT 2004 THOMSON DERWENT on STN
TI New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia
AN AAV44747 cDNA DGENE
AB This sequence represents an expressed sequence tag, which is specifically stated as not being contained within the DNA of the invention. The DNA of the invention encodes the human oncogene induced secreted protein I (**HOIPS I**). **HOIPS I** can be used for treating cell proliferative diseases, particularly cancers, such as acute and chronic myelogenous leukaemias. The products can also be used for detection and diagnosis of a cell proliferative or cell differentiation disorders.
ACCESSION NUMBER: AAV44747 cDNA DGENE
TITLE: New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia
INVENTOR: Olsen H S; Ruben S M

PATENT ASSIGNEE: (HUMA-N)HUMAN GENOME SCI INC.
PATENT INFO: WO 9828421 A1 19980702 71p
APPLICATION INFO: WO 1997-US23547 19971219
PRIORITY INFO: US 1997-37388 19970207
US 1996-33869 19961220
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 1998-377652 [32]
DESCRIPTION: Expressed sequence tag AA340310.

L3 ANSWER 17 OF 18 DGENE COPYRIGHT 2004 THOMSON DERWENT on STN
TI New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia
AN AAV44745 cDNA DGENE
AB This sequence encodes the human oncogene induced secreted protein I (**HOIPS I**) of the invention. **HOIPS I** can be used for treating cell proliferative diseases, particularly cancers, such as acute and chronic myelogenous leukaemias. The products can also be used for detection and diagnosis of a cell proliferative or cell differentiation disorders.
ACCESSION NUMBER: AAV44745 cDNA DGENE
TITLE: New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia
INVENTOR: Olsen H S; Ruben S M
PATENT ASSIGNEE: (HUMA-N)HUMAN GENOME SCI INC.
PATENT INFO: WO 9828421 A1 19980702 71p
APPLICATION INFO: WO 1997-US23547 19971219
PRIORITY INFO: US 1997-37388 19970207
US 1996-33869 19961220
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 1998-377652 [32]
CROSS REFERENCES: P-PSDB: AAW69220
DESCRIPTION: Human oncogene induced secreted protein I coding sequence.

L3 ANSWER 18 OF 18 WPIDS COPYRIGHT 2004 THOMSON DERWENT on STN
TI New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia.
AN 1998-377652 [32] WPIDS
AB WO 9828421 A UPAB: 19980812
An isolated nucleic acid molecule (I) is claimed comprising a polynucleotide (PN) having a nucleotide sequence (NS) at least 95% identical to a sequence selected from:

(a) a NS encoding a polypeptide comprising amino acids from -20 to 142, -19 to 142, or 1 to 142 of the 162 amino acid (aa) sequence given in the specification (sequence representing a Human Oncogene Induced Secreted Protein I (**HOIPS I**) polypeptide);

(b) a NS encoding a polypeptide having an amino acid sequence encoded by a cDNA clone contained in ATCC 97825;

(c) a NS encoding a mature **HOIPS I** polypeptide having an amino acid sequence encoded by a cDNA clone contained in ATCC 97825; and

(d) a NS complementary to any of the NSs in (a)-(c).

Also claimed are:

(1) an isolated Nucleic Acid Molecule (NAM) comprising a PN which hybridises under stringent hybridisation conditions (I) where the PN which hybridises does not hybridise under stringent hybridisation conditions to a PN having a NS consisting of only A residues or of only T residues;

(2) an isolated NAM comprising a PN which encodes an amino acid sequence of an epitope-bearing portion of a **HOIPS I**

polypeptide having an amino acid sequence as in (a)-(c) above;
(3) an isolated NAM comprising a PN having a sequence at least 95% identical to a sequence selected from:
(a) a NS of a fragment of a 860 bp sequence given in the specification (encoding the **HOIPS I** polypeptide), where the fragment comprises at least 50 contiguous nucleotides of the 860 bp, provided that the NS is not one of the 514, 457, 413, 320, 264, and 249 sequences given in the specification; and
(b) a NS complementary to a NS as in (a);
(4) a method for making a recombinant vector comprising inserting (I) into a vector;
(5) a recombinant vector produced by a method as in (4);
(6) a method of making a recombinant host cell comprising introducing a recombinant vector as in (5) into a host cell;
(7) a recombinant host cell produced by a method as in (6);
(8) an isolated **HOIPS I** polypeptide having an amino acid sequence at least 95% identical to a sequence encoded by (I) or an epitope-bearing portion of the polypeptide;
(9) an isolated polypeptide comprising an epitope-bearing portion of the **HOIPS I** protein, where the portion is selected from a polypeptide comprising amino acid residues from -4 to 9, from 13 to 19, from 23 to 32, from 36 to 47, from 54 to 63, from 70 to 74, from 90 to 100, from 105 to 119 or from 125 to 132 of the 162 aa sequence;
(10) an isolated **HOIPS I** polypeptide where, except for 1 to 50 conservative amino acid substitutions, the polypeptide has a sequence selected from:
(a) amino acids from -20 to 142, 19 to 142, or 1 to 142 of the 162 aa sequence given in the specification;
(b) an amino acid sequence of the **HOIPS I** polypeptide having an amino acid sequence encoded by a cDNA contained in ATCC 97825;
(c) an amino acid sequence of a mature **HOIPS I** polypeptide having an amino acid sequence encoded by a cDNA clone contained in ATCC 97825; and
(d) an amino acid sequence of an epitope-bearing portion of any one of the polypeptides as in (a)-(c);
(11) an isolated nucleic acid encoding a polypeptide as in (10).
USE - The products can be used for treating cell proliferative diseases, particularly cancers, such as acute and chronic myelogenous leukaemias. The products can also be used for detection and diagnosis.
Dwg.0/3

ACCESSION NUMBER: 1998-377652 [32] WPIDS
DOC. NO. NON-CPI: N1998-295209
DOC. NO. CPI: C1998-114764
TITLE: New isolated human oncogene induced secreted protein - used to develop products for the diagnosis and treatment of cell proliferative diseases, particularly cancers such as leukaemia.
DERWENT CLASS: B04 D16 S03
INVENTOR(S): OLSEN, H S; RUBEN, S M
PATENT ASSIGNEE(S): (HUMA-N) HUMAN GENOME SCI INC
COUNTRY COUNT: 81
PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
WO 9828421	A1	19980702	(199832)*	EN	71
RW: AT BE CH DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC MW NL OA PT SD SE SZ UG ZW					
W: AL AM AT AU AZ BA BB BG BR BY CA CH CN CU CZ DE DK EE ES FI GB GE GH GM GW HU ID IL IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MD MG MK MN MW MX NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT UA UG US UZ VN YU ZW					
AU 9858027	A	19980717	(199848)		

US 6284486 B1 20010904 (200154)
US 2002119552 A1 20020829 (200259)

APPLICATION DETAILS:

PATENT NO	KIND		APPLICATION	DATE
WO 9828421	A1		WO 1997-US23547	19971219
AU 9858027	A		AU 1998-58027	19971219
US 6284486	B1	Provisional	US 1996-33869P	19961220
		Provisional	US 1997-37388P	19970207
			US 1997-994962	19971219
US 2002119552	A1	Provisional	US 1996-33869P	19961220
		Provisional	US 1997-37388P	19970207
		Div ex	US 1997-994962	19971219
			US 2001-899917	20010709

FILING DETAILS:

PATENT NO	KIND		PATENT NO
AU 9858027	A	Based on	WO 9828421

PRIORITY APPLN. INFO: US 1997-37388P 19970207; US
1996-33869P 19961220; US
1997-994962 19971219; US
2001-899917 20010709

Refine Search

Search Results -

Terms	Documents
L9 and L10	18

Database:
US Pre-Grant Publication Full-Text Database
US Patents Full-Text Database
US OCR Full-Text Database
EPO Abstracts Database
JPO Abstracts Database
Derwent World Patents Index
IBM Technical Disclosure Bulletins

Search: L11

Refine Search

Recall Text Clear Interrupt

Search History

DATE: Thursday, April 29, 2004 Printable Copy Create Case

Set Name side by side	Query	Hit Count	Set Name result set
DB=USPT; PLUR=YES; OP=OR			
L11	l9 and L10	18	L11
L10	ruben.in.	1520	L10
L9	l6 and L8	31	L9
L8	olsen.in.	2539	L8
L7	henrik-olsen.in.	0	L7
L6	myelogenous leukemia and L5	10350	L6
L5	myeloid proliferation and l4	29144	L5
L4	L3 and protein sequence	614907	L4
L3	human oncogene induced secreted protein I	1403401	L3
L2	HOIPs-I	0	L2
L1	6284486.pn.	1	L1

END OF SEARCH HISTORY

h e b b cg b e e ch

Hit List

Clear | Generate Collection | Print | Fwd Refs | Bkwd Refs | Generate OACS

Search Results - Record(s) 1 through 10 of 18 returned.

☐ 1. Document ID: US 6632920 B1

L11: Entry 1 of 18 File: USPT Oct 14, 2003

US-PAT-NO: 6632920
DOCUMENT-IDENTIFIER: US 6632920 B1

TITLE: 36 human secreted proteins

DATE-ISSUED: October 14, 2003

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Olsen; Henrik S.	Gaithersburg	MD		
Ruben; Steven M.	Olney	MD		
Rosen; Craig A.	Laytonsville	MD		
Brewer; Laurie A.	St. Paul	MN		
Ebner; Reinhard	Gaithersburg	MD		
Duan; Roxanne	Bethesda	MD		
Florence; Kimberly	Rockville	MD		

US-CL-CURRENT: 530/300; 435/69.1, 530/324

Full | Title | Citation | Front | Review | Classification | Date | Reference | Sequences | Attachments | Claims | KWIC | Draw De

☐ 2. Document ID: US 6627741 B2

L11: Entry 2 of 18 File: USPT Sep 30, 2003

US-PAT-NO: 6627741
DOCUMENT-IDENTIFIER: US 6627741 B2

TITLE: Antibodies to secreted protein HCEJQ69

DATE-ISSUED: September 30, 2003

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Ruben; Steven M.	Olney	MD		
Ni; Jian	Germantown	MD		

h e b b g e e e f e ef b e

Rosen; Craig A.	Laytonsville	MD
Wei; Ying-Fei	Berkeley	CA
Young; Paul	Gaithersburg	MD
Florence; Kimberly	Rockville	MD
Soppet; Daniel R.	Centreville	VA
Brewer; Laurie A.	St. Paul	MN
Endress; Gregory A.	Florence	MA
Carter; Kenneth C.	North Potomac	MD
Mucenski; Michael	Cincinnati	OH
Ebner; Reinhard	Gaithersburg	MD
LaFleur; David W.	Washington	DC
Olsen; Henrik	Gaithersburg	MD
Shi; Yanggu	Gaithersburg	MD
Moore; Paul A.	Germantown	MD
Komatsoulis; George	Silver Spring	MD

US-CL-CURRENT: 530/389.2; 530/387.1, 530/387.3, 530/387.7, 530/387.9, 530/388.1, 530/388.15, 530/389.1

Full | Title | Citation | Front | Review | Classification | Date | Reference | Claims | KWIC | Draw De

---

☐ 3. Document ID: US 6605592 B2

L11: Entry 3 of 18 File: USPT Aug 12, 2003

US-PAT-NO: 6605592
DOCUMENT-IDENTIFIER: US 6605592 B2

TITLE: Protein HOFNF53

DATE-ISSUED: August 12, 2003

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Ni; Jian	Germantown	MD		
Baker; Kevin P.	Darnestown	MD		
Birse; Charles E.	North Potomac	MD		
Ebner; Reinhard	Gaithersburg	MD		
Fiscella; Michele	Bethesda	MD		
Komatsoulis; George A.	Silver Spring	MD		
LaFleur; David W.	Washington	DC		
Moore; Paul A.	Germantown	MD		
Olsen; Henrik S.	Gaithersburg	MD		
Rosen; Craig A.	Laytonsville	MD		
Ruben; Steven M.	Olney	MD		
Soppet; Daniel R.	Centreville	VA		
Young; Paul E.	Gaithersburg	MD		
Wei; Ping	Brookeville	MD		

h e b b g e c e f e ef b e

Florence; Kimberly A. Rockville MD

US-CL-CURRENT: 514/2; 435/252.3, 435/254.11, 435/320.1, 435/325, 435/471, 435/69.1, 435/71.1, 435/71.2, 514/12, 514/8, 530/350

Full | Title | Citation | Front | Review | Classification | Date | Reference | Claims | KWMC | Draw De

---

☐ 4. Document ID: US 6590075 B2

L11: Entry 4 of 18 File: USPT Jul 8, 2003

US-PAT-NO: 6590075
DOCUMENT-IDENTIFIER: US 6590075 B2

TITLE: Secreted protein HODAZ50

DATE-ISSUED: July 8, 2003

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Ruben; Steven M.	Olney	MD		
Rosen; Craig A.	Laytonsville	MD		
Fischer; Carrie L.	Burke	VA		
Soppet; Daniel R.	Centreville	VA		
Carter; Kenneth C.	North Potomac	MD		
Bednarik; Daniel P.	Columbia	MD		
Endress; Gregory A.	Potomac	MD		
Yu; Guo-Liang	Berkeley	CA		
Ni; Jian	Rockville	MD		
Feng; Ping	Gaithersburg	MD		
Young; Paul E.	Gaithersburg	MD		
Greene; John M.	Gaithersburg	MD		
Ferrie; Ann M.	Tewksbury	MA		
Duan; Roxanne	Bethesda	MD		
Hu; Jing-Shan	Sunnyvale	CA		
Florence; Kimberly A.	Rockville	MD		
Olsen; Henrik S.	Gaithersburg	MD		
Ebner; Reinhard	Gaithersburg	MD		
Brewer; Laurie A.	St. Paul	MN		
Shi; Yanggu	Gaithersburg	MD		

US-CL-CURRENT: 530/350; 435/6, 435/69.1, 435/7.1, 536/23.1

Full | Title | Citation | Front | Review | Classification | Date | Reference | Claims | KWMC | Draw De

---

☐ 5. Document ID: US 6576445 B1

L11: Entry 5 of 18 File: USPT Jun 10, 2003

h e b b g c e e f e ef b e

US-PAT-NO: 6576445
DOCUMENT-IDENTIFIER: US 6576445 B1

TITLE: Chemokine .alpha.-4

DATE-ISSUED: June 10, 2003

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Olsen; Henrik S.	Gaithersburg	MD		
Ruben; Steven M.	Olney	MD		
Zeng; Zhi-Zhen	Gaithersburg	MD		

US-CL-CURRENT: 435/69.5; 435/325, 530/300, 530/350, 530/386, 530/387.1, 530/387.3, 530/387.9, 530/388.1, 530/388.15, 530/388.23, 530/389.1, 530/389.2, 530/391.1, 530/391.3, 530/391.5, 530/391.7, 530/391.9

Full | Title | Citation | Front | Review | Classification | Date | Reference | Claims | KWIC | Draw De

---

☐ 6. Document ID: US 6566325 B2

L11: Entry 6 of 18 File: USPT May 20, 2003

US-PAT-NO: 6566325
DOCUMENT-IDENTIFIER: US 6566325 B2

TITLE: 49 human secreted proteins

DATE-ISSUED: May 20, 2003

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Moore; Paul A.	Germantown	MD		
Ruben; Steven M.	Olney	MD		
Olsen; Henrik S.	Gaithersburg	MD		
Shi; Yanggu	Gaithersburg	MD		
Rosen; Craig A.	Laytonsville	MD		
Florence; Kimberly A.	Rockville	MD		
Soppet; Daniel R.	Centreville	VA		
LaFleur; David W.	Washington	DC		
Endress; Gregory A.	Potomac	MD		
Ebner; Reinhard	Gaithersburg	MD		
Komatsoulis; George	Silver Spring	MD		
Duan; Roxanne D.	Bethesda	MD		

US-CL-CURRENT: 514/2; 530/300, 530/350

Full | Title | Citation | Front | Review | Classification | Date | Reference | Claims | KWIC | Draw De

---

h e b b g e e e f e ef b e

☐ 7. Document ID: US 6548063 B1

L11: Entry 7 of 18 File: USPT Apr 15, 2003

US-PAT-NO: 6548063
DOCUMENT-IDENTIFIER: US 6548063 B1

TITLE: Synferon

DATE-ISSUED: April 15, 2003

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Olsen; Henrik S.	Gaithersburg	MD		
Gentz; Reiner L.	Rockville	MD		
Ruben; Steven M.	Olney	MD		

US-CL-CURRENT: 424/158.1; 424/85.4, 435/325, 530/350, 530/351, 530/388.1, 530/389.1

Full | Title | Citation | Front | Review | Classification | Date | Reference | Claims | KWMC | Draw De

---

☐ 8. Document ID: US 6534631 B1

L11: Entry 8 of 18 File: USPT Mar 18, 2003

US-PAT-NO: 6534631
DOCUMENT-IDENTIFIER: US 6534631 B1

TITLE: Secreted protein HT5GJ57

DATE-ISSUED: March 18, 2003

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Ruben; Steven M.	Olney	MD		
Komatsoulis; George	Silver Spring	MD		
Duan; Roxanne D.	Bethesda	MD		
Rosen; Craig A.	Laytonsville	MD		
Moore; Paul A.	Germantown	MD		
Shi; Yanggu	Gaithersburg	MD		
LaFleur; David W.	Washington	DC		
Ebner; Reinhard	Gaithersburg	MD		
Olsen; Henrik	Gaithersburg	MD		
Brewer; Laurie A.	St. Paul	MN		
Florence; Kimberly A.	Rockville	MD		
Young; Paul	Gaithersburg	MD		
Mucenski; Michael	Cincinnati	OH		
Endress; Gregory A.	Potomac	MD		
Soppet; Daniel R.	Centreville	VA		

h e b b g e e e f e ef b e

US-CL-CURRENT: 530/350; 435/320.1, 435/325, 530/300, 536/23.1, 536/24.1

Full | Title | Citation | Front | Review | Classification | Date | Reference | Sequences | Attachments | Claims | KWIC | Draw De

☐ 9. Document ID: US 6525174 B1

L11: Entry 9 of 18 File: USPT Feb 25, 2003

US-PAT-NO: 6525174
DOCUMENT-IDENTIFIER: US 6525174 B1

TITLE: Precerebellin-like protein

DATE-ISSUED: February 25, 2003

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Young; Paul	Gaithersburg	MD		
Greene; John M.	Gaithersburg	MD		
Ferrie; Ann M.	Tewksbury	MA		
Ruben; Steven M.	Olney	MD		
Rosen; Craig A.	Laytonsville	MD		
Hu; Jing-Shan	Sunnyvale	CA		
Olsen; Henrik S.	Gaithersburg	MD		
Ebner; Reinhard	Gaithersburg	MD		
Brewer; Laurie A.	St. Paul	MN		
Moore; Paul A.	Germantown	MD		
Shi; Yanggu	Gaithersburg	MD		
Florence; Charles	Rockville	MD		
Florence; Kimberly	Rockville	MD		
Lafleur; David W.	Washington	DC		
Ni; Jian	Rockville	MD		
Fan; Ping	Gaithersburg	MD		
Wei; Ying-Fei	Berkeley	CA		
Fischer; Carrie L.	Burke	VA		
Soppet; Daniel R.	Centreville	VA		
Li; Yi	Sunnyvale	CA		
Zeng; Zhizhen	Gaithersburg	MD		
Kyaw; Hla	Frederick	MD		
Yu; Guo-Liang	Berkeley	CA		
Feng; Ping	Gaithersburg	MD		
Dillon; Patrick J.	Carlsbad	CA		
Endress; Gregory A.	Potomac	MD		
Carter; Kenneth C.	North Potomac	MD		

US-CL-CURRENT: 530/350; 435/69.1

Full | Title | Citation | Front | Review | Classification | Date | Reference | Sequences | Attachments | Claims | KWIC | Draw De

h e b b g e e e f e ef b e

☐ 10. Document ID: US 6476195 B1

L11: Entry 10 of 18 File: USPT Nov 5, 2002

US-PAT-NO: 6476195
DOCUMENT-IDENTIFIER: US 6476195 B1

TITLE: Secreted protein HNFGF20

DATE-ISSUED: November 5, 2002

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Komatsoulis; George	Silver Spring	MD		
Rosen; Craig A.	Laytonsville	MD		
Ruben; Steven M.	Olney	MD		
Duan; Roxanne D.	Bethesda	MD		
Moore; Paul A.	Germantown	MD		
Shi; Yanggu	Gaithersburg	MD		
LaFleur; David W.	Washington	DC		
Wei; Ying-Fei	Berkeley	CA		
Ni; Jian	Rockville	MD		
Florence; Kimberly A.	Rockville	MD		
Young; Paul	Gaithersburg	MD		
Brewer; Laurie A.	St. Paul	MN		
Soppet; Daniel R.	Centreville	VA		
Endress; Gregory A.	Potomac	MD		
Ebner; Reinhard	Gaithersburg	MD		
Olsen; Henrik	Gaithersburg	MD		
Mucenski; Michael	Cincinnati	OH		

US-CL-CURRENT: 530/350; 435/6, 435/7.1, 536/23.1

Full | Title | Citation | Front | Review | Classification | Date | Reference | Sequences | Attachments | Claims | KWIC | Draw De

Clear | Generate Collection | Print | Fwd Refs | Bkwd Refs | Generate OACS

Terms	Documents
L9 and L10	18

Display Format: CIT Change Format

Previous Page Next Page Go to Doc#

h e b b g e e e f e ef b e

Search Forms
Search Results
Help
User Searches
Preferences
Logout

Hit List

Clear | Generate Collection | Print | Fwd Refs | Bkwd Refs | Generate OACS

Search Results - Record(s) 11 through 18 of 18 returned.

☐ 11. Document ID: US 6475753 B1

L11: Entry 11 of 18 File: USPT Nov 5, 2002

US-PAT-NO: 6475753
DOCUMENT-IDENTIFIER: US 6475753 B1

TITLE: 94 Human Secreted Proteins

DATE-ISSUED: November 5, 2002

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Ruben; Steven M.	Olney	MD		
Ni; Jian	Rockville	MD		
Rosen; Craig A.	Laytonsville	MD		
Wei; Ying-Fei	Berkeley	CA		
Young; Paul	Gaithersburg	MD		
Florence; Kimberly	Rockville	MD		
Soppet; Daniel R.	Centreville	VA		
Brewer; Laurie A.	St. Paul	MN		
Endress; Gregory A.	Potomac	MD		
Carter; Kenneth C.	Potomac	MD		
Mucenski; Michael	Cincinnati	OH		
Ebner; Reinhard	Gaithersburg	MD		
Lafleur; David W.	Washington	DC		
Olsen; Henrik	Gaithersburg	MD		
Shi; Yanggu	Gaithersburg	MD		
Moore; Paul A.	Germantown	MD		
Komatsoulis; George	Silver Spring	MD		

US-CL-CURRENT: 435/69.1; 435/252.3, 435/320.1, 435/325, 435/471, 435/69.4, 435/71.1, 530/350, 536/23.5

Full | Title | Citation | Front | Review | Classification | Date | Reference | Sequences | Attachments | Claims | KWMC | Draw De

☐ 12. Document ID: US 6433139 B1

L11: Entry 12 of 18 File: USPT Aug 13, 2002

h e b b cg b cc e

US-PAT-NO: 6433139
DOCUMENT-IDENTIFIER: US 6433139 B1

TITLE: Secreted protein HPEAD48

DATE-ISSUED: August 13, 2002

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Ruben; Steven M.	Olney	MD		
Rosen; Craig A.	Laytonsville	MD		
Olsen; Henrik S.	Gaithersburg	MD		

US-CL-CURRENT: 530/350; 435/6, 435/7.1, 530/300, 536/23.1

Full | Title | Citation | Front | Review | Classification | Date | Reference | Claims | KWIC | Draw De

---

☐ 13. Document ID: US 6420526 B1

L11: Entry 13 of 18 File: USPT Jul 16, 2002

US-PAT-NO: 6420526
DOCUMENT-IDENTIFIER: US 6420526 B1

TITLE: 186 human secreted proteins

DATE-ISSUED: July 16, 2002

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Ruben; Steven M.	Olney	MD		
Rosen; Craig A.	Laytonsville	MD		
Fischer; Carrie L.	Burke	VA		
Soppet; Daniel P.	Centreville	VA		
Carter; Kenneth C.	North Potomac	MD		
Bednarik; Daniel R.	Columbia	MD		
Endress; Gregory A.	Potomac	MD		
Yu; Guo-Liang	Berkeley	CA		
Ni; Jian	Rockville	MD		
Feng; Ping	Gaithersburg	MD		
Young; Paul E.	Gaithersburg	MD		
Greene; John M.	Gaithersburg	MD		
Ferrie; Ann M.	Tewksbury	MA		
Duan; Roxanne	Bethesda	MD		
Hu; Jing-Shan	Sunnyvale	CA		
Florence; Kimberly A.	Rockville	MD		
Olsen; Henrik S.	Gaithersburg	MD		
Ebner; Reinhard	Gaithersburg	MD		
Brewer; Laurie A.	St. Paul	MN		

h e b b cg b cc e

Moore; Paul A.	Germantown	MD
Shi; Yanggu	Gaithersburg	MD
Lafleur; David W.	Washington	DC
Li; Yi	Sunnyvale	CA
Zeng; Zhizhen	Lansdale	PA
Kyaw; Hla	Frederick	MD

US-CL-CURRENT: 530/350; 435/6, 536/23.1

Full | Title | Citation | Front | Review | Classification | Date | Reference | Sequences | Attachments | Claims | KWMC | Draw. De

---

☐ 14. Document ID: US 6379926 B1

L11: Entry 14 of 18 File: USPT Apr 30, 2002

US-PAT-NO: 6379926
DOCUMENT-IDENTIFIER: US 6379926 B1

TITLE: Polynucleotides encoding chemokine .beta.-6 antagonists

DATE-ISSUED: April 30, 2002

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Kreider; Brent L.	Germantown	MD		
Ruben; Steven M.	Olney	MD		
Olsen; Henrik S.	Gaithersburg	MD		

US-CL-CURRENT: 435/69.5; 435/252.3, 435/320.1, 435/325, 435/455, 435/471, 435/69.1, 435/69.7, 536/23.1, 536/23.4, 536/23.5

Full | Title | Citation | Front | Review | Classification | Date | Reference | Sequences | Attachments | Claims | KWMC | Draw. De

---

☐ 15. Document ID: US 6342581 B1

L11: Entry 15 of 18 File: USPT Jan 29, 2002

US-PAT-NO: 6342581
DOCUMENT-IDENTIFIER: US 6342581 B1
**** See image for Certificate of Correction ****

TITLE: Secreted protein HLHFP03

DATE-ISSUED: January 29, 2002

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Rosen; Craig A.	Laytonsville	MD		
Ruben; Steven M.	Olney	MD		

h e b b cg b cc e

Olsen; Henrik S.	Gaithersburg	MD
Ebner; Reinhard	Gaithersburg	MD

US-CL-CURRENT: 530/300; 435/69.1, 530/350

Full | Title | Citation | Front | Review | Classification | Date | Reference | References | Attachments | Claims | KWMC | Draw De

---

☐ 16. Document ID: US 6284486 B1

L11: Entry 16 of 18 File: USPT Sep 4, 2001

US-PAT-NO: 6284486
DOCUMENT-IDENTIFIER: US 6284486 B1

TITLE: Human oncogene induced secreted protein I

DATE-ISSUED: September 4, 2001

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Olsen; Henrik S.	Gaithersburg	MD		
Ruben; Steven M.	Olney	MD		

US-CL-CURRENT: 435/69.1; 435/252.3, 435/6, 435/91.1, 435/91.4, 435/91.41, 536/23.1, 536/23.5, 536/24.1, 536/24.2

Full | Title | Citation | Front | Review | Classification | Date | Reference | References | Attachments | Claims | KWMC | Draw De

---

☐ 17. Document ID: US 6114145 A

L11: Entry 17 of 18 File: USPT Sep 5, 2000

US-PAT-NO: 6114145
DOCUMENT-IDENTIFIER: US 6114145 A

TITLE: Synferon, a synthetic interferon

DATE-ISSUED: September 5, 2000

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Olsen; Henrik S.	Gaithersburg	MD		
Gentz; Reiner L.	Rockville	MD		
Ruben; Steven M.	Olney	MD		

US-CL-CURRENT: 435/69.51; 424/85.4, 435/252.3, 435/252.33, 435/254.11, 435/325, 435/455, 435/471, 435/91.41, 530/351, 536/23.52

Full | Title | Citation | Front | Review | Classification | Date | Reference | References | Attachments | Claims | KWMC | Draw De

h e b b cg b cc e

☐ 18. Document ID: US 5916769 A

L11: Entry 18 of 18 File: USPT Jun 29, 1999

US-PAT-NO: 5916769
DOCUMENT-IDENTIFIER: US 5916769 A

TITLE: Polynucleotides encoding extra cellular/epidermal growth factor HCABA58X polypepides

DATE-ISSUED: June 29, 1999

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Olsen; Henrik S.	Gaithersburg	MD		
Ruben; Steven M.	Olney	MD		

US-CL-CURRENT: 435/69.4; 435/243, 435/320.1, 435/325, 435/69.1, 536/23.1, 536/23.51

Full | Title | Citation | Front | Review | Classification | Date | Reference | Claims | KWIC | Draw De

Clear | Generate Collection | Print | Fwd Refs | Bkwd Refs | Generate OACS

Terms	Documents
L9 and L10	18

Display Format: CIT Change Format

Previous Page Next Page Go to Doc#

h e b b cg b cc e

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: April 29, 2004, 15:21:06 ; Search time 58 Seconds
(without alignments)
789.184 Million cell updates/sec

Title: US-09-899-917-2
Perfect score: 862
Sequence: 1 MKGFTATLFLWTLIFPSCSG.....ELYTEKRSTVACANATIMCS 162

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	862	100.0	162	2	AAW69220	Aaw69220 Human onc
2	862	100.0	162	2	AAY13933	Aay13933 Human OHP
3	862	100.0	162	7	ADE59274	Ade59274 Human Pro
4	854	99.1	162	3	AAY77476	Aay77476 Human MD-
5	850	98.6	162	6	ABU61841	Abu61841 Human MD1
6	758	87.9	162	3	AAY84000	Aay84000 Human B-c
7	613	71.1	162	2	AAY13932	Aay13932 Mouse OHP
8	613	71.1	162	3	AAY77481	Aay77481 Rodent pr
9	613	71.1	162	6	ABU61846	Abu61846 Mouse MD1
10	613	71.1	162	7	ADE59272	Ade59272 Rat Prote
11	336	39.0	160	3	AAY77482	Aay77482 Avian pro
12	336	39.0	160	6	ABU61847	Abu61847 Chicken M
13	106	12.3	160	2	AAY13929	Aay13929 Mouse OHP
14	106	12.3	160	3	AAB35928	Aab35928 Murine MD
15	101.5	11.8	160	2	AAY13931	Aay13931 Human OHP
16	101.5	11.8	160	3	AAB35927	Aab35927 Human MD-
17	101.5	11.8	160	3	AAY77478	Aay77478 Human MD-
18	101.5	11.8	160	6	ABU61843	Abu61843 Updated h
19	96.5	11.2	141	3	AAY77480	Aay77480 Primate M
20	96.5	11.2	141	6	ABU61845	Abu61845 Mouse MD2
21	79.5	9.2	484	5	ABB91964	Abb91964 Herbicida
22	78	9.0	876	6	ABU49402	Abu49402 Protein e
23	77	8.9	363	3	AAG28947	Aag28947 Arabidops
24	77	8.9	364	3	AAG28946	Aag28946 Arabidops
25	76.5	8.9	1177	6	ABU17986	Abu17986 Protein e
26	75.5	8.8	342	6	ABU26566	Abu26566 Protein e
27	75.5	8.8	1200	3	AAG30768	Aag30768 Arabidops
28	75.5	8.8	1207	3	AAG30767	Aag30767 Arabidops
29	75.5	8.8	1215	3	AAG30766	Aag30766 Arabidops
30	75.5	8.8	1241	6	ABU24033	Abu24033 Protein e
31	75	8.7	377	2	AAY32037	Aay32037 Bovine pr
32	75	8.7	505	7	ADB75603	Adb75603 Prostate
33	75	8.7	505	7	ADE38437	Ade38437 Human pro
34	75	8.7	524	3	AAB56960	Aab56960 Human pro
35	75	8.7	699	5	ABB91291	Abb91291 Herbicida
36	74	8.6	392	2	AAY32057	Aay32057 Bovine pr
37	73.5	8.5	264	2	AAR29640	Aar29640 pCTD ORF
38	72.5	8.4	851	7	ADC64335	Adc64335 KIAA1196.
39	72.5	8.4	1172	5	ABP63009	Abp63009 Human pol
40	72.5	8.4	1178	5	ABP63008	Abp63008 Human pol
41	72	8.4	269	4	AAB95041	Aab95041 Human pro
42	72	8.4	608	5	ABP65007	Abp65007 Human pro
43	72	8.4	700	5	ABB92360	Abb92360 Herbicida
44	72	8.4	790	5	AAO19184	Aao19184 Human neu
45	71.5	8.3	310	6	ABM70549	Abm70549 Photorhab

ALIGNMENTS

RESULT 1
AAW69220
ID AAW69220 standard; protein; 162 AA.
XX
AC AAW69220;
XX
DT 16-OCT-1998 (first entry)
XX
DE Human oncogene induced secreted protein I.
XX
KW Oncogene induced secreted protein I; HOIPS I; human; cancer; therapy;
KW cell proliferative disease; myelogenous leukaemia;
KW cell differentiation disorder.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1. .20
FT /note= "signal peptide"
FT Protein 21. .162
FT /note= "mature HOIPS I"
XX
PN WO9828421-A1.
XX
PD 02-JUL-1998.
XX
PF 19-DEC-1997; 97WO-US023547.
XX
PR 20-DEC-1996; 96US-0033869P.
PR 07-FEB-1997; 97US-0037388P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Olsen HS, Ruben SM;
XX
DR WPI; 1998-377652/32.
DR N-PSDB; AAV44745.
XX
PT New isolated human oncogene induced secreted protein - used to develop
PT products for the diagnosis and treatment of cell proliferative diseases,
PT particularly cancers such as leukaemia.
XX
PS Claim 12; Page 51-52; 71pp; English.
XX
CC This sequence is the human oncogene induced secreted protein I (HOIPS I)
CC of the invention. HOIPS I can be used for treating cell proliferative
CC diseases, particularly cancers, such as acute and chronic myelogenous
CC leukaemias. The products can also be used for detection and diagnosis of

```
CC   a cell proliferative or cell differentiation disorders
XX
SQ   Sequence 162 AA;

  Query Match           100.0%;  Score 862;  DB 2;  Length 162;
  Best Local Similarity  100.0%;  Pred. No. 9e-92;
  Matches  162;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MKGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSK 60

Qy         61 QLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 QLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQI 120

Qy        121 YYAGPVNNPEFTIPQGEYQVLLELYTEKRSTVACANATIMCS 162
              ||||||||||||||||||||||||||||||||||||||||||
Db        121 YYAGPVNNPEFTIPQGEYQVLLELYTEKRSTVACANATIMCS 162


RESULT 2
AAY13933
ID   AAY13933 standard; protein; 162 AA.
XX
AC   AAY13933;
XX
DT   14-JUL-1999  (first entry)
XX
DE   Human OHP106 protein.
XX
KW   OHP106; bacterial; viral; infection; cytokine activity; tissue formation;
KW   cell proliferation; cell differentiation; immunostimulation; therapy;
KW   immunosuppression; haematopoiesis control; tissue repair.
XX
OS   Homo sapiens.
XX
PN   WO9918205-A1.
XX
PD   15-APR-1999.
XX
PF   06-OCT-1998;  98WO-JP004515.
XX
PR   07-OCT-1997;  97JP-00274673.
XX
PA   (ONOY ) ONO PHARM CO LTD.
XX
PI   Honjo T,  Kato K,  Tada H;
XX
DR   WPI; 1999-277269/23.
DR   N-PSDB; AAX36695, AAX36696.
XX
PT   New polypeptides useful in the treatment of various diseases including
PT   bacterial and viral infections with cytokine activity.
XX
PS   Claim 1; Page 47; 57pp; Japanese.
XX
CC   This sequence represents a OHP106 protein of the invention. The
CC   polypeptides and their antibodies are applicable in drug compositions,
CC   useful for treatment of various diseases including bacterial and viral
CC   infections with cytokine activity, and activities e.g. on cell
CC   proliferation/differentiation, immunostimulation/suppression, control of
CC   haematopoiesis, and tissue formation and repair. The proteins and
CC   antibodies have a wide spectrum of activities
XX
SQ   Sequence 162 AA;

  Query Match           100.0%;  Score 862;  DB 2;  Length 162;
  Best Local Similarity  100.0%;  Pred. No. 9e-92;
  Matches  162;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MKGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSK 60

Qy         61 QLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 QLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQI 120

Qy        121 YYAGPVNNPEFTIPQGEYQVLLELYTEKRSTVACANATIMCS 162
              ||||||||||||||||||||||||||||||||||||||||||
Db        121 YYAGPVNNPEFTIPQGEYQVLLELYTEKRSTVACANATIMCS 162


RESULT 3
ADE59274
ID   ADE59274 standard; protein; 162 AA.
XX
AC   ADE59274;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human Protein O95711, SEQ ID NO 5168.
XX
KW   Human; pain; neuronal tissue; gene therapy;
KW   spinal segmental nerve injury; chronic constriction injury; CCI;
KW   spared nerve injury; SNI; Chung.
XX
OS   Homo sapiens.
XX
PN   WO2003016475-A2.
XX
PD   27-FEB-2003.
XX
PF   14-AUG-2002; 2002WO-US025765.
XX
PR   14-AUG-2001; 2001US-0312147P.
PR   01-NOV-2001; 2001US-0346382P.
PR   26-NOV-2001; 2001US-0333347P.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
PA   (FARB ) BAYER AG.
XX
PI   Woolf C,  D'urso D,  Befort K,  Costigan M;
XX
DR   WPI; 2003-268312/26.
DR   GENBANK; O95711.
XX
PT   New composition comprising two or more isolated polypeptides, useful for
PT   preparing a medicament for treating pain in an animal.
XX
PS   Claim 1; Page; 1017pp; English.
XX
CC   The invention discloses a composition comprising two or more isolated rat
CC   or human polynucleotides or a polynucleotide which represents a fragment,
CC   derivative or allelic variation of the nucleic acid sequence. Also
CC   claimed are a vector comprising the novel polynucleotide, a host cell
CC   comprising the vector, a method for identifying a nucleotide sequence
CC   which is differentially regulated in an animal subjected to pain and a
CC   kit to perform the method, an array, a method for identifying an agent
CC   that increases or decreases the expression of the polynucleotide sequence
CC   that is differentially expressed in neuronal tissue of a first animal
CC   subjected to pain, a method for identifying a compound which regulates
CC   the expression of a polynucleotide sequence which is differentially
CC   expressed in an animal subjected to pain, a method for identifying a
CC   compound that regulates the activity of one or more of the
CC   polynucleotides, a method for producing a pharmaceutical composition, a
CC   method for identifying a compound or small molecule that regulates the
CC   activity in an animal of one or more of the polypeptides given in the
CC   specification, a method for identifying a compound useful in treating
CC   pain and a pharmaceutical composition comprising the one or more
CC   polypeptides or their antibodies. The polynucleotide or the compound that
CC   modulates its activity is useful for preparing a medicament for treating
CC   pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC   injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC   therapy). The sequence presented is a human protein (shown in Table 2 of
```

bad date

```
CC   the specification) which is differentially expressed during pain. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic form directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 162 AA;

  Query Match             100.0%;  Score 862;  DB 7;  Length 162;
  Best Local Similarity   100.0%;  Pred. No. 9e-92;
  Matches  162;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MKGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSK 60

Qy         61 QLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 QLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQI 120

Qy        121 YYAGPVNNPEFTIPQGEYQVLLELYTEKRSTVACANATIMCS 162
              |||||||||||||||||||||||||||||||||||||||||
Db        121 YYAGPVNNPEFTIPQGEYQVLLELYTEKRSTVACANATIMCS 162


RESULT 4
AAY77476
ID   AAY77476 standard; protein; 162 AA.
XX
AC   AAY77476;
XX
DT   05-JUN-2000  (first entry)
XX
DE   Human MD-1 protein.
XX
KW   MD-1; MD-2; leucine-rich motif; LRR protein ligand; immune disorder;
KW   inflammation; allergy; immunosuppressant; antiarthritic; antirheumatoid;
KW   antiinflammatory; dermatological; antithyroid.
XX
OS   Homo sapiens.
XX
PN   WO200001817-A2.
XX
PD   13-JAN-2000.
XX
PF   06-JUL-1999;   99WO-US012366.
XX
PR   06-JUL-1998;   98US-00110938.
PR   13-JUL-1998;   98US-00114466.
PR   23-JUL-1998;   98US-0093897P.
PR   12-AUG-1998;   98US-00132968.
PR   18-AUG-1998;   98US-00136214.
PR   11-SEP-1998;   98US-0099999P.
XX
PA   (SCHE ) SCHERING CORP.
XX
PI   Bates EEM,  Lebecque SJE,  Murphy EE,  Mattson JD,  Gorman DM;
PI   Hedrick JA,  Wang L,  Zlotnik A,  Murgolo NJ,  Greene JR,  Johnston JA;
PI   Bazan JF,  Mahony D,  Lees EM;
XX
DR   WPI; 2000-171015/15.
DR   N-PSDB; AAZ92417.
XX
PT   New isolated mammalian genes, used to develop products for treating e.g.
PT   immune, inflammatory or allergic abnormalities, cancers or degenerative
PT   conditions.
XX
PS   Claim 43; Page 203-204; 218pp; English.
XX
CC   The invention relates to a number of primate and/or rodent proteins, and
CC   the genes which encode them. The invention encompasses human dendritic
CC   cell prostaglandin transporter (DC-PGT); the TNF (tumour necrosis factor)
CC   receptor family-related proteins HDTEA84, HSLJD37R and RANKL; human CC
CC   chemokine HCC5; human deubiquitinating proteins Dub11 and Dub 12; human
CC   MD-1 and human and murine MD-2 proteins, which exhibit the properties of
CC   ligands for proteins comprising a leucine-rich motif (LRR); human cyclin
CC   E2; cDNAs encoding these proteins; and antibodies against these proteins.
CC   The proteins can be used for modulating the physiology or development of
CC   a cell. They can be used to mediate uptake of substrates (e.g.,
CC   prostaglandin-like molecules), to modulate or mediate cellular
CC   interactions (e.g., induce or prevent trafficking, proliferation, or
CC   differentiation of cells), or are intracellular proteins which are
CC   important in various cellular processes such as the deubiquitination of
CC   proteins or cell cycle regulation. The products can be used for treating
CC   medical conditions such as immune, inflammatory or allergic disorders, or
CC   abnormal cellular proliferation, for example, cancers or degenerative
CC   conditions. They can be used to modulate immune responses in disease
CC   states e.g., autoimmune disorders, including rheumatoid arthritis,
CC   systemic lupus erythematosus, Hashimoto's autoimmune thyroiditis, as well
CC   as acute and chronic inflammatory responses in which T cell activation,
CC   expansion, and/or immunological T cell memory play an important role.
CC   Sequences AAY77476-Y77480 represent MD-1 and MD-2 proteins. AAY77476 is
CC   human MD-1, AAY77477-Y77478 are human MD-2 proteins, AAY77479 is murine
CC   MD-2 and AAY77480 is a primate MD-2
XX
SQ   Sequence 162 AA;

  Query Match             99.1%;  Score 854;  DB 3;  Length 162;
  Best Local Similarity   99.4%;  Pred. No. 7.7e-91;
  Matches  161;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MKGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSK 60
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db          1 MKGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSRLEVLYQSCDPLQDFGFSVEKCSK 60

Qy         61 QLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 QLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQI 120

Qy        121 YYAGPVNNPEFTIPQGEYQVLLELYTEKRSTVACANATIMCS 162
              |||||||||||||||||||||||||||||||||||||||||
Db        121 YYAGPVNNPEFTIPQGEYQVLLELYTEKRSTVACANATIMCS 162


RESULT 5
ABU61841
ID   ABU61841 standard; protein; 162 AA.
XX
AC   ABU61841;
XX
DT   15-AUG-2003  (first entry)
XX
DE   Human MD1.
XX
KW   Human; MD1; cell development; tissue culture development.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 13
FT                   /label= Unknown
FT                   /note= "Encoded by CTM"
FT   Misc-difference 37
FT                   /label= Unknown
FT                   /note= "Encoded by SKY"
XX
PN   US2003028005-A1.
XX
PD   06-FEB-2003.
XX
PF   11-AUG-1999;   99US-00372348.
XX
PR   12-AUG-1998;   98US-0096328P.
XX
PA   (BAZA/) BAZAN J F.
XX
PI   Bazan JF;
```

```
XX
DR   WPI; 2003-466159/44.
DR   N-PSDB; ACA62461.
XX
PT   New composition comprising HCC5, Dub11, Dub12, primate MD1 or MD2 or
PT   rodent MD2 polypeptide, useful for modulating physiology or development
PT   of a cell or tissue culture cells.
XX
PS   Claim 1; Page 15-16; 77pp; English.
XX
CC   The invention relates to a novel chemokine, de-ubiquitination or cell
CC   surface protein. The composition is useful for modulating physiology or
CC   development of a cell or tissue culture cells. The present sequence
CC   represents the amino acid sequence of human MD1
XX
SQ   Sequence 162 AA;

  Query Match             98.6%;  Score 850;  DB 6;  Length 162;
  Best Local Similarity   98.8%;  Pred. No. 2.2e-90;
  Matches  160;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MKGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSK 60
              ||||||||||| |||||||||||||||||||| |||||||||||||||||||||||||||
Db          1 MKGFTATLFLWTXIFPSCSGGGGGKAWPTHVVCSDSXLEVLYQSCDPLQDFGFSVEKCSK 60

Qy         61 QLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 QLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQI 120

Qy        121 YYAGPVNNPEFTIPQGEYQVLLELYTEKRSTVACANATIMCS 162
              |||||||||||||||||||||||||||||||||||||||||
Db        121 YYAGPVNNPEFTIPQGEYQVLLELYTEKRSTVACANATIMCS 162


RESULT 6
AAY84000
ID   AAY84000 standard; protein; 162 AA.
XX
AC   AAY84000;
XX
DT   25-OCT-2000  (first entry)
XX
DE   Human B-cell surface protein MD-1.
XX
KW   Human; B cell surface protein; apoptosis; antiallergic; allergy; EST;
KW   antiasthmatic; immunosuppressive; dermatological; antiinflammatory;
KW   autoimmune disease; asthma; atopic dermatitis; expressed sequence tag.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .19
FT                   /note= "signal peptide"
FT   Protein         20. .162
FT                   /note= "mature MD-1 protein"
XX
PN   JP2000106880-A.
XX
PD   18-APR-2000.
XX
PF   08-OCT-1998;   98JP-00286470.
XX
PR   08-OCT-1998;   98JP-00286470.
XX
PA   (MITU ) MITSUBISHI CHEM CORP.
XX
DR   WPI; 2000-353476/31.
DR   N-PSDB; AAA11221.
XX
PT   A new DNA encoding a human B cell surface protein, useful for treating
PT   autoimmune diseases, allergic diseases, asthma and atopic dermatitis.
XX
PS   Claim 1; Page 11; 12pp; Japanese.
XX
CC   This sequence represents a human B cell surface protein, designated MD-1.
CC   The protein is involved in the inhibition or induction of B cell
CC   apoptosis and has antiallergic, antiasthmatic, immunosuppressive,
CC   dermatological and antiinflammatory activity. The protein and the DNA are
CC   used as treating and diagnostic agents for diseases related to antibody
CC   production by B cells, e.g. autoimmune or allergic diseases, asthma and
CC   atopic dermatitis. The coding sequence was isolated from a human EST
CC   clone (GenBank accession number T84854)
XX
SQ   Sequence 162 AA;

  Query Match             87.9%;  Score 758;  DB 3;  Length 162;
  Best Local Similarity   88.9%;  Pred. No. 1.1e-79;
  Matches  144;  Conservative    3;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          1 MKGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSK 60

Qy         61 QLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQI 120
              |||||||||||| |   :     :  | :||||||||||||||||||||||||||||||
Db         61 QLKSNINIRFGIISVEKCSKQLKSNINIRFGITVLNFSYPICEAALPKFSFCGRRKGEQI 120

Qy        121 YYAGPVNNPEFTIPQGEYQVLLELYTEKRSTVACANATIMCS 162
              |||||||||||||||||||||||||||||||||||||||||
Db        121 YYAGPVNNPEFTIPQGEYQVLLELYTEKRSTVACANATIMCS 162


RESULT 7
AAY13932
ID   AAY13932 standard; protein; 162 AA.
XX
AC   AAY13932;
XX
DT   14-JUL-1999  (first entry)
XX
DE   Mouse OHP106 protein.
XX
KW   OHP106; bacterial; viral; infection; cytokine activity; tissue formation;
KW   cell proliferation; cell differentiation; immunostimulation; therapy;
KW   immunosuppression; haematopoiesis control; tissue repair.
XX
OS   Mus musculus.
XX
PN   WO9918205-A1.
XX
PD   15-APR-1999.
XX
PF   06-OCT-1998;   98WO-JP004515.
XX
PR   07-OCT-1997;   97JP-00274673.
XX
PA   (ONOY ) ONO PHARM CO LTD.
XX
PI   Honjo T,  Kato K,  Tada H;
XX
DR   WPI; 1999-277269/23.
DR   N-PSDB; AAX36693, AAX36694.
XX
PT   New polypeptides useful in the treatment of various diseases including
PT   bacterial and viral infections with cytokine activity.
XX
PS   Claim 1; Page 44; 57pp; Japanese.
XX
CC   This sequence represents a OHP106 protein of the invention. The
CC   polypeptides and their antibodies are applicable in drug compositions,
CC   useful for treatment of various diseases including bacterial and viral
CC   infections with cytokine activity, and activities e.g. on cell
CC   proliferation/differentiation, immunostimulation/suppression, control of
CC   haematopoiesis, and tissue formation and repair. The proteins and
CC   antibodies have a wide spectrum of activities
XX
```

```
SQ   Sequence 162 AA;

  Query Match             71.1%;  Score 613;  DB 2;  Length 162;
  Best Local Similarity   66.7%;  Pred. No. 8.5e-63;
  Matches  108;  Conservative   22;  Mismatches   32;  Indels    0;  Gaps    0;

Qy          1 MKGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSK 60
              |  |  | | :| |  || |  |    || |  |:  ||||:|||||||||||| |:::|||
Db          1 MNGVAAALLVWILTSPSSSDHGSENGWPKHTACNSGGLEVVYQSCDPLQDFGLSIDQCSK 60

Qy         61 QLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQI 120
              |::||:|||||||||:||::||||: ||::|||:||:|||:||   ||||||||||||||
Db         61 QIQSNLNIRFGIILRQDIRKLFLDITLMAKGSSILNYSYPLCEEDQPKFSFCGRRKGEQI 120

Qy        121 YYAGPVNNPEFTIPQGEYQVLLELYTEKRSTVACANATIMCS 162
              |||||||||   :||||||:||||| | |:||||||||:  |
Db        121 YYAGPVNNPGLDVPQGEYQLLLELYNENRATVACANATVTSS 162


RESULT 8
AAY77481
ID   AAY77481 standard; protein; 162 AA.
XX
AC   AAY77481;
XX
DT   05-JUN-2000  (first entry)
XX
DE   Rodent protein sequence, SEQ ID NO:50.
XX
KW   Imune disorder; inflammation; allergy; immunosuppressant; antiarthritic;
KW   antirheumatoid; antiinflammatory; dermatological; antithyroid.
XX
OS   Rodentia.
XX
PN   WO200001817-A2.
XX
PD   13-JAN-2000.
XX
PF   06-JUL-1999;   99WO-US012366.
XX
PR   06-JUL-1998;   98US-00110938.
PR   13-JUL-1998;   98US-00114466.
PR   23-JUL-1998;   98US-0093897P.
PR   12-AUG-1998;   98US-00132968.
PR   18-AUG-1998;   98US-00136214.
PR   11-SEP-1998;   98US-0099999P.
XX
PA   (SCHE ) SCHERING CORP.
XX
PI   Bates EEM,  Lebecque SJE,  Murphy EE,  Mattson JD,  Gorman DM;
PI   Hedrick JA,  Wang L,  Zlotnik A,  Murgolo NJ,  Greene JR,  Johnston JA;
PI   Bazan JF,  Mahony D,  Lees EM;
XX
DR   WPI; 2000-171015/15.
XX
PT   New isolated mammalian genes, used to develop products for treating e.g.
PT   immune, inflammatory or allergic abnormalities, cancers or degenerative
PT   conditions.
XX
PS   Disclosure; Page 208-209; 218pp; English.
XX
CC   The invention relates to a number of primate and/or rodent proteins, and
CC   the genes which encode them. The invention encompasses human dendritic
CC   cell prostaglandin transporter (DC-PGT); the TNF (tumour necrosis factor)
CC   receptor family-related proteins HDTEA84, HSLJD37R and RANKL; human CC
CC   chemokine HCC5; human deubiquitinating proteins Dub11 and Dub 12; human
CC   MD-1 and human and murine MD-2 proteins, which exhibit the properties of
CC   ligands for proteins comprising a leucine-rich motif (LRR); human cyclin
CC   E2; cDNAs encoding these proteins; and antibodies against these proteins.
CC   The proteins can be used for modulating the physiology or development of
CC   a cell. They can be used to mediate uptake of substrates (e.g.,
CC   prostaglandin-like molecules), to modulate or mediate cellular
CC   interactions (e.g., induce or prevent trafficking, proliferation, or
CC   differentiation of cells), or are intracellular proteins which are
CC   important in various cellular processes such as the deubiquitination of
CC   proteins or cell cycle regulation. The products can be used for treating
CC   medical conditions such as immune, inflammatory or allergic disorders, or
CC   abnormal cellular proliferation, for example, cancers or degenerative
CC   conditions. They can be used to modulate immune responses in disease
CC   states e.g., autoimmune disorders, including rheumatoid arthritis,
CC   systemic lupus erythematosus, Hashimoto's autoimmune thyroiditis, as well
CC   as acute and chronic inflammatory responses in which T cell activation,
CC   expansion, and/or immunological T cell memory play an important role.
CC   Sequences AAY77463-Y77464, AAY77474-Y77475 and AAY77484 represent primate
CC   proteins of undefined function, AAY77462 and AAY77481 are rodent proteins
CC   of undefined function, and AAY77482 is an avian protein of undefined
CC   function. These sequences are given in the sequence listing but are not
CC   referred to in the remainder of the specification
XX
SQ   Sequence 162 AA;

  Query Match             71.1%;  Score 613;  DB 3;  Length 162;
  Best Local Similarity   66.7%;  Pred. No. 8.5e-63;
  Matches  108;  Conservative   22;  Mismatches   32;  Indels    0;  Gaps    0;

Qy          1 MKGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSK 60
              |  |  | | :| |  || |  |    || |  |:  ||||:|||||||||||| |:::|||
Db          1 MNGVAAALLVWILTSPSSSDHGSENGWPKHTACNSGGLEVVYQSCDPLQDFGLSIDQCSK 60

Qy         61 QLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQI 120
              |::||:|||||||||:||::||||: ||::|||:||:|||:||   ||||||||||||||
Db         61 QIQSNLNIRFGIILRQDIRKLFLDITLMAKGSSILNYSYPLCEEDQPKFSFCGRRKGEQI 120

Qy        121 YYAGPVNNPEFTIPQGEYQVLLELYTEKRSTVACANATIMCS 162
              |||||||||   :||||||:||||| | |:||||||||:  |
Db        121 YYAGPVNNPGLDVPQGEYQLLLELYNENRATVACANATVTSS 162


RESULT 9
ABU61846
ID   ABU61846 standard; protein; 162 AA.
XX
AC   ABU61846;
XX
DT   15-AUG-2003  (first entry)
XX
DE   Mouse MD1.
XX
KW   Mouse; MD1; cell development; tissue culture development.
XX
OS   Mus sp.
XX
PN   US2003028005-A1.
XX
PD   06-FEB-2003.
XX
PF   11-AUG-1999;   99US-00372348.
XX
PR   12-AUG-1998;   98US-0096328P.
XX
PA   (BAZA/) BAZAN J F.
XX
PI   Bazan JF;
XX
DR   WPI; 2003-466159/44.
XX
PT   New composition comprising HCC5, Dub11, Dub12, primate MD1 or MD2 or
PT   rodent MD2 polypeptide, useful for modulating physiology or development
PT   of a cell or tissue culture cells.
XX
PS   Disclosure; Page 18; 77pp; English.
XX
CC   The invention relates to a novel chemokine, de-ubiquitination or cell
CC   surface protein. The composition is useful for modulating physiology or
CC   development of a cell or tissue culture cells. The present sequence
CC   represents the amino acid sequence of mouse MD1
```

```
XX
SQ   Sequence 162 AA;

  Query Match             71.1%;  Score 613;  DB 6;  Length 162;
  Best Local Similarity   66.7%;  Pred. No. 8.5e-63;
  Matches  108;  Conservative   22;  Mismatches   32;  Indels    0;  Gaps    0;

Qy          1 MKGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSK 60
              | |  | | :| |  || |  |    || |  |:  ||||:|||||||||||| |:::|||
Db          1 MNGVAAALLVWILTSPSSSDHGSENGWPKHTACNSGGLEVVYQSCDPLQDFGLSIDQCSK 60

Qy         61 QLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQI 120
              |::||:|||||||||:||::||||: ||::|||:||:|||:||   ||||||||||||||
Db         61 QIQSNLNIRFGIILRQDIRKLFLDITLMAKGSSILNYSYPLCEEDQPKFSFCGRRKGEQI 120

Qy        121 YYAGPVNNPEFTIPQGEYQVLLELYTEKRSTVACANATIMCS 162
              |||||||||   :||||||:|||||  | |:||||||||:  |
Db        121 YYAGPVNNPGLDVPQGEYQLLLELYNENRATVACANATVTSS 162


RESULT 10
ADE59272
ID   ADE59272 standard; protein; 162 AA.
XX
AC   ADE59272;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Rat Protein NP_034875, SEQ ID NO 5166.
XX
KW   Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW   chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS   Rattus norvegicus.
XX
PN   WO2003016475-A2.
XX
PD   27-FEB-2003.
XX
PF   14-AUG-2002; 2002WO-US025765.
XX
PR   14-AUG-2001; 2001US-0312147P.
PR   01-NOV-2001; 2001US-0346382P.
PR   26-NOV-2001; 2001US-0333347P.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
PA   (FARB ) BAYER AG.
XX
PI   Woolf C,  D'urso D,  Befort K,  Costigan M;
XX
DR   WPI; 2003-268312/26.
DR   GENBANK; NP_034875.
XX
PT   New composition comprising two or more isolated polypeptides, useful for
PT   preparing a medicament for treating pain in an animal.
XX
PS   Claim 1; Page; 1017pp; English.
XX
CC   The invention discloses a composition comprising two or more isolated rat
CC   or human polynucleotides or a polynucleotide which represents a fragment,
CC   derivative or allelic variation of the nucleic acid sequence. Also
CC   claimed are a vector comprising the novel polynucleotide, a host cell
CC   comprising the vector, a method for identifying a nucleotide sequence
CC   which is differentially regulated in an animal subjected to pain and a
CC   kit to perform the method, an array, a method for identifying an agent
CC   that increases or decreases the expression of the polynucleotide sequence
CC   that is differentially expressed in neuronal tissue of a first animal
CC   subjected to pain, a method for identifying a compound which regulates
CC   the expression of a polynucleotide sequence which is differentially
CC   expressed in an animal subjected to pain, a method for identifying a
CC   compound that regulates the activity of one or more of the
CC   polynucleotides, a method for producing a pharmaceutical composition, a
CC   method for identifying a compound or small molecule that regulates the
CC   activity in an animal of one or more of the polypeptides given in the
CC   specification, a method for identifying a compound useful in treating
CC   pain and a pharmaceutical composition comprising the one or more
CC   polypeptides or their antibodies. The polynucleotide or the compound that
CC   modulates its activity is useful for preparing a medicament for treating
CC   pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC   injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC   therapy). The sequence presented is a rat protein (shown in Table 2 of
CC   the specification) which is differentially expressed during pain. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic form directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 162 AA;

  Query Match             71.1%;  Score 613;  DB 7;  Length 162;
  Best Local Similarity   66.7%;  Pred. No. 8.5e-63;
  Matches  108;  Conservative   22;  Mismatches   32;  Indels    0;  Gaps    0;

Qy          1 MKGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSK 60
              | |  | | :| |  || |  |    || |  |:  ||||:|||||||||||| |:::|||
Db          1 MNGVAAALLVWILTSPSSSDHGSENGWPKHTACNSGGLEVVYQSCDPLQDFGLSIDQCSK 60

Qy         61 QLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQI 120
              |::||:|||||||||:||::||||: ||::|||:||:|||:||   ||||||||||||||
Db         61 QIQSNLNIRFGIILRQDIRKLFLDITLMAKGSSILNYSYPLCEEDQPKFSFCGRRKGEQI 120

Qy        121 YYAGPVNNPEFTIPQGEYQVLLELYTEKRSTVACANATIMCS 162
              |||||||||   :||||||:|||||  | |:||||||||:  |
Db        121 YYAGPVNNPGLDVPQGEYQLLLELYNENRATVACANATVTSS 162


RESULT 11
AAY77482
ID   AAY77482 standard; protein; 160 AA.
XX
AC   AAY77482;
XX
DT   05-JUN-2000  (first entry)
XX
DE   Avian protein sequence, SEQ ID NO:52.
XX
KW   Imune disorder; inflammation; allergy; immunosuppressant; antiarthritic;
KW   antirheumatoid; antiinflammatory; dermatological; antithyroid.
XX
OS   Aves.
XX
PN   WO200001817-A2.
XX
PD   13-JAN-2000.
XX
PF   06-JUL-1999;   99WO-US012366.
XX
PR   06-JUL-1998;   98US-00110938.
PR   13-JUL-1998;   98US-00114466.
PR   23-JUL-1998;   98US-0093897P.
PR   12-AUG-1998;   98US-00132968.
PR   18-AUG-1998;   98US-00136214.
PR   11-SEP-1998;   98US-0099999P.
XX
PA   (SCHE ) SCHERING CORP.
XX
PI   Bates EEM,  Lebecque SJE,  Murphy EE,  Mattson JD,  Gorman DM;
PI   Hedrick JA,  Wang L,  Zlotnik A,  Murgolo NJ,  Greene JR,  Johnston JA;
PI   Bazan JF,  Mahony D,  Lees EM;
XX
DR   WPI; 2000-171015/15.
DR   N-PSDB; AAZ92421.
XX
PT   New isolated mammalian genes, used to develop products for treating e.g.
PT   immune, inflammatory or allergic abnormalities, cancers or degenerative
PT   conditions.
XX
```

PS Disclosure; Page 210-211; 218pp; English.
XX
CC The invention relates to a number of primate and/or rodent proteins, and
CC the genes which encode them. The invention encompasses human dendritic
CC cell prostaglandin transporter (DC-PGT); the TNF (tumour necrosis factor)
CC receptor family-related proteins HDTEA84, HSLJD37R and RANKL; human CC
CC chemokine HCC5; human deubiquitinating proteins Dub11 and Dub 12; human
CC MD-1 and human and murine MD-2 proteins, which exhibit the properties of
CC ligands for proteins comprising a leucine-rich motif (LRR); human cyclin
CC E2; cDNAs encoding these proteins; and antibodies against these proteins.
CC The proteins can be used for modulating the physiology or development of
CC a cell. They can be used to mediate uptake of substrates (e.g.,
CC prostaglandin-like molecules), to modulate or mediate cellular
CC interactions (e:g., induce or prevent trafficking, proliferation, or
CC differentiation of cells), or are intracellular proteins which are
CC important in various cellular processes such as the deubiquitination of
CC proteins or cell cycle regulation. The products can be used for treating
CC medical conditions such as immune, inflammatory or allergic disorders, or
CC abnormal cellular proliferation, for example, cancers or degenerative
CC conditions. They can be used to modulate immune responses in disease
CC states e.g., autoimmune disorders, including rheumatoid arthritis,
CC systemic lupus erythematosus, Hashimoto's autoimmune thyroiditis, as well
CC as acute and chronic inflammatory responses in which T cell activation,
CC expansion, and/or immunological T cell memory play an important role.
CC Sequences AAY77463-Y77464, AAY77474-Y77475 and AAY77484 represent primate
CC proteins of undefined function, AAY77462 and AAY77481 are rodent proteins
CC of undefined function, and AAY77482 is an avian protein of undefined
CC function. These sequences are given in the sequence listing but are not
CC referred to in the remainder of the specification
XX
SQ Sequence 160 AA;

Query Match 39.0%; Score 336; DB 3; Length 160;
Best Local Similarity 44.4%; Pred. No. 1.4e-30;
Matches 59; Conservative 26; Mismatches 48; Indels 0; Gaps 0;

```
Qy         27 WPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSKQLKSNINIRFGIILREDIKELFLDLA 86
              |||| || :  ||: |:|||| ||| ||:::||       :||  ::||: ||||:  :
Db         22 WPTHTVCKEENLEIYYKSCDPQQDFAFSIDRCSDVTTHTFDIRAAMVLRQSIKELYAKVD 81

Qy         87 LMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQIYYAGPVNNPEFTIPQGEYQVLLELYT 146
              |:  | :||::|  :|   | |  |||::||| :|| ||:      ||| :| :   |
Db         82 LIINGKTVLSYSETLCGPGLSKLIFCGKKKGEHLYYEGPITLGIKEIPQRDYTITARLTN 141

Qy        147 EKRSTVACANATI 159
              | |:|||||: |:
Db        142 EDRATVACADFTV 154
```

RESULT 12
ABU61847
ID ABU61847 standard; protein; 160 AA.
XX
AC ABU61847;
XX
DT 15-AUG-2003 (first entry)
XX
DE Chicken MD1.
XX
KW Chicken; MD1; cell development; tissue culture development.
XX
OS Gallus gallus.
XX
PN US2003028005-A1.
XX
PD 06-FEB-2003.
XX
PF 11-AUG-1999; 99US-00372348.
XX
PR 12-AUG-1998; 98US-0096328P.
XX
PA (BAZA/) BAZAN J F.
XX
PI Bazan JF;
XX
DR WPI; 2003-466159/44.
DR N-PSDB; ACA62465.
XX
PT New composition comprising HCC5, Dub11, Dub12, primate MD1 or MD2 or
PT rodent MD2 polypeptide, useful for modulating physiology or development
PT of a cell or tissue culture cells.
XX
PS Disclosure; Page 18; 77pp; English.
XX
CC The invention relates to a novel chemokine, de-ubiquitination or cell
CC surface protein. The composition is useful for modulating physiology or
CC development of a cell or tissue culture cells. The present sequence
CC represents the amino acid sequence of chicken MD1
XX
SQ Sequence 160 AA;

Query Match 39.0%; Score 336; DB 6; Length 160;
Best Local Similarity 44.4%; Pred. No. 1.4e-30;
Matches 59; Conservative 26; Mismatches 48; Indels 0; Gaps 0;

```
Qy         27 WPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSKQLKSNINIRFGIILREDIKELFLDLA 86
              |||| || :  ||: |:|||| ||| ||:::||       :||  ::||: ||||:  :
Db         22 WPTHTVCKEENLEIYYKSCDPQQDFAFSIDRCSDVTTHTFDIRAAMVLRQSIKELYAKVD 81

Qy         87 LMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQIYYAGPVNNPEFTIPQGEYQVLLELYT 146
              |:  | :||::|  :|   | |  |||::||| :|| ||:      ||| :| :   |
Db         82 LIINGKTVLSYSETLCGPGLSKLIFCGKKKGEHLYYEGPITLGIKEIPQRDYTITARLTN 141

Qy        147 EKRSTVACANATI 159
              | |:|||||: |:
Db        142 EDRATVACADFTV 154
```

RESULT 13
AAY13929
ID AAY13929 standard; protein; 160 AA.
XX
AC AAY13929;
XX
DT 14-JUL-1999 (first entry)
XX
DE Mouse OHP106 protein.
XX
KW OHP106; bacterial; viral; infection; cytokine activity; tissue formation;
KW cell proliferation; cell differentiation; immunostimulation; therapy;
KW immunosuppression; haematopoiesis control; tissue repair.
XX
OS Mus musculus.
XX
PN WO9918205-A1.
XX
PD 15-APR-1999.
XX
PF 06-OCT-1998; 98WO-JP004515.
XX
PR 07-OCT-1997; 97JP-00274673.
XX
PA (ONOY) ONO PHARM CO LTD.
XX
PI Honjo T, Kato K, Tada H;
XX
DR WPI; 1999-277269/23.
DR N-PSDB; AAX36687, AAX36688.
XX
PT New polypeptides useful in the treatment of various diseases including
PT bacterial and viral infections with cytokine activity.
XX
PS Claim 1; Page 36-37; 57pp; Japanese.
XX
CC This sequence represents a OHP106 protein of the invention. The
CC polypeptides and their antibodies are applicable in drug compositions,

```
CC   useful for treatment of various diseases including bacterial and viral
CC   infections with cytokine activity, and activities e.g. on cell
CC   proliferation/differentiation, immunostimulation/suppression, control of
CC   haematopoiesis, and tissue formation and repair. The proteins and
CC   antibodies have a wide spectrum of activities
XX
SQ   Sequence 160 AA;

  Query Match             12.3%;  Score 106;  DB 2;  Length 160;
  Best Local Similarity   26.3%;  Pred. No. 0.00084;
  Matches   36;  Conservative   24;  Mismatches   61;  Indels   16;  Gaps    5;

Qy         33 CSDSGLEVLYQSCDPLQ-DFGFSVEKCSKQLKSNINIRFGIILREDIKELFLDLALMSQG 91
              |: |   : |  || |:     | | | :   :|  :    | | ::| |: :| :
Db         25 CNSSDAIISYSYCDHLKFPISISSEPCIRLRGTNGFVHVEFIPRGNLKYLYFNLFI---- 80

Qy         92 SSVLNFSYP-----ICEAALPKFSFCGRRKGEQIYYAGPVNNPEFTIPQGEYQVLLELY- 145
               || :   |     :|      :|||   ||| :  : | :      :| |: : |
Db         81 -SVNSIELPKRKEVLCHGHDDDYSFCRALKGETVNTSIPFSFEGILFPKGHYRCVAEAIA 139

Qy        146 --TEKRSTVACANATIM 160
                ||::  : | | ||:
Db        140 GDTEEK--LFCLNFTII 154


RESULT 14
AAB35928
ID   AAB35928 standard; protein; 160 AA.
XX
AC   AAB35928;
XX
DT   26-FEB-2001  (first entry)
XX
DE   Murine MD-2 amino acid sequence.
XX
KW   MD-2; immunosuppressant; allergy; asthma; atopic dermatitis; mouse;
KW   infection; prevention.
XX
OS   Mus musculus.
XX
PN   JP2000262290-A.
XX
PD   26-SEP-2000.
XX
PF   18-MAR-1999;   99JP-00073815.
XX
PR   18-MAR-1999;   99JP-00073815.
XX
PA   (MITU ) MITSUBISHI CHEM CORP.
XX
DR   WPI; 2000-682118/67.
DR   N-PSDB; AAC66388.
XX
PT   A new protein controlling the infection-preventing reaction against a
PT   microbe.
XX
PS   Claim 2; Page 11-13; 14pp; Japanese.
XX
CC   This invention relates to a protein, MD-2, which controls an infection
CC   prevention reaction against a microbe. Protein and cDNA sequences
CC   AAC66387 - AAC66388 and AAB35927 - AAB35928 represent human and mouse MD-
CC   2. MD-2 increases NFKappaB activation through TLP4 molecules. The protein
CC   has immunosuppressant, antiallergy; antiasthmatic and dermatological
CC   activity. The protein and the DNA can be used as a treating agent and a
CC   diagnostic agent for diseases such as autoimmune diseases, allergic
CC   diseases, asthma and atopic dermatitis
XX
SQ   Sequence 160 AA;

  Query Match             12.3%;  Score 106;  DB 3;  Length 160;
  Best Local Similarity   26.3%;  Pred. No. 0.00084;
  Matches   36;  Conservative   24;  Mismatches   61;  Indels   16;  Gaps    5;

Qy         33 CSDSGLEVLYQSCDPLQ-DFGFSVEKCSKQLKSNINIRFGIILREDIKELFLDLALMSQG 91
              |: |   : |  || |:     | | | :   :|  :    | | ::| |: :| :
Db         25 CNSSDAIISYSYCDHLKFPISISSEPCIRLRGTNGFVHVEFIPRGNLKYLYFNLFI---- 80

Qy         92 SSVLNFSYP-----ICEAALPKFSFCGRRKGEQIYYAGPVNNPEFTIPQGEYQVLLELY- 145
               || :   |     :|      :|||   ||| :  : | :      :| |: : |
Db         81 -SVNSIELPKRKEVLCHGHDDDYSFCRALKGETVNTSIPFSFEGILFPKGHYRCVAEAIA 139

Qy        146 --TEKRSTVACANATIM 160
                ||::  : | | ||:
Db        140 GDTEEK--LFCLNFTII 154


RESULT 15
AAY13931
ID   AAY13931 standard; protein; 160 AA.
XX
AC   AAY13931;
XX
DT   14-JUL-1999  (first entry)
XX
DE   Human OHP106 protein.
XX
KW   OHP106; bacterial; viral; infection; cytokine activity; tissue formation;
KW   cell proliferation; cell differentiation; immunostimulation; therapy;
KW   immunosuppression; haematopoiesis control; tissue repair.
XX
OS   Homo sapiens.
XX
PN   WO9918205-A1.
XX
PD   15-APR-1999.
XX
PF   06-OCT-1998;   98WO-JP004515.
XX
PR   07-OCT-1997;   97JP-00274673.
XX
PA   (ONOY ) ONO PHARM CO LTD.
XX
PI   Honjo T,  Kato K,  Tada H;
XX
DR   WPI; 1999-277269/23.
DR   N-PSDB; AAX36691, AAX36692.
XX
PT   New polypeptides useful in the treatment of various diseases including
PT   bacterial and viral infections with cytokine activity.
XX
PS   Claim 1; Page 41-42; 57pp; Japanese.
XX
CC   This sequence represents a OHP106 protein of the invention. The
CC   polypeptides and their antibodies are applicable in drug compositions,
CC   useful for treatment of various diseases including bacterial and viral
CC   infections with cytokine activity, and activities e.g. on cell
CC   proliferation/differentiation, immunostimulation/suppression, control of
CC   haematopoiesis, and tissue formation and repair. The proteins and
CC   antibodies have a wide spectrum of activities
XX
SQ   Sequence 160 AA;

  Query Match             11.8%;  Score 101.5;  DB 2;  Length 160;
  Best Local Similarity   24.1%;  Pred. No. 0.0028;
  Matches   34;  Conservative   26;  Mismatches   70;  Indels   11;  Gaps    3;

Qy          4 FTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQ-DFGFSVEKCSKQL 62
              | :|||  : ||         :|   : ||: |   : |  || :|     :|  | :
Db          6 FFSTLF--SSIFT--------EAQKQYWVCNSSDASISYTYCDKMQYPISINVNPCIELK 55

Qy         63 KSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQIYY 122
               |   :    | | |:|:|: :| :     ::       ||  :   :|||   ||| :
Db         56 GSKGLLHIFYIPRRDLKQLYFNLYITVNTMNLPKRKEVICRGSDDDYSFCRALKGETVNT 115

Qy        123 AGPVNNPEFTIPQGEYQVLLE 143
                    :      :|:|: ::|
```

Db 116 TISFSFKGIKFSKGKYKCVVE 136

Search completed: April 29, 2004, 15:22:20
Job time : 60 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: April 29, 2004, 15:21:06 ; Search time 23 Seconds
(without alignments)
363.626 Million cell updates/sec

Title: US-09-899-917-2
Perfect score: 862
Sequence: 1 MKGFTATLFLWTLIFPSCSG.....ELYTEKRSTVACANATIMCS 162

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	862	100.0	162	3	US-08-994-962-2	Sequence 2, Appli
2	344	39.9	133	3	US-08-994-962-3	Sequence 3, Appli
3	94	10.9	60	3	US-08-994-962-17	Sequence 17, Appl
4	73.5	8.5	264	3	US-08-969-644-16	Sequence 16, Appl
5	73.5	8.5	264	3	US-08-444-189-16	Sequence 16, Appl
6	73.5	8.5	264	3	US-08-468-544-16	Sequence 16, Appl
7	73.5	8.5	442	1	US-08-220-151-22	Sequence 22, Appl
8	73.5	8.5	442	1	US-08-413-118-22	Sequence 22, Appl
9	73.5	8.5	442	3	US-08-473-446-22	Sequence 22, Appl
10	73.5	8.5	442	6	5470718-3	Patent No. 5470718
11	73	8.5	787	4	US-09-252-991A-26468	Sequence 26468, A
12	71.5	8.3	442	6	5470718-2	Patent No. 5470718
13	71	8.2	668	1	US-08-530-950-13	Sequence 13, Appl
14	71	8.2	668	3	US-09-149-879-13	Sequence 13, Appl
15	71	8.2	668	4	US-09-057-009-13	Sequence 13, Appl
16	69.5	8.1	184	4	US-09-384-162-12	Sequence 12, Appl
17	69	8.0	1012	4	US-09-285-385C-4	Sequence 4, Appli
18	68	7.9	294	4	US-09-252-991A-27242	Sequence 27242, A
19	68	7.9	459	1	US-08-336-618-12	Sequence 12, Appl
20	68	7.9	459	1	US-08-336-618-26	Sequence 26, Appl
21	68	7.9	857	1	US-07-717-331F-2	Sequence 2, Appli
22	67.5	7.8	204	4	US-09-134-001C-3617	Sequence 3617, Ap
23	67.5	7.8	458	1	US-08-336-618-24	Sequence 24, Appl
24	67.5	7.8	844	4	US-09-564-805-227	Sequence 227, App
25	66.5	7.7	372	1	US-08-597-236-8	Sequence 8, Appli
26	66.5	7.7	372	1	US-08-746-682A-8	Sequence 8, Appli
27	66.5	7.7	793	2	US-08-468-558-5	Sequence 5, Appli
28	66.5	7.7	793	3	US-08-676-444-5	Sequence 5, Appli
29	66	7.7	494	3	US-09-205-283-12	Sequence 12, Appl
30	66	7.7	2233	2	US-08-569-853-1	Sequence 1, Appli
31	66	7.7	2233	2	US-08-569-853-2	Sequence 2, Appli
32	66	7.7	2233	3	US-08-987-439-1	Sequence 1, Appli
33	65.5	7.6	617	4	US-09-198-452A-155	Sequence 155, App
34	65	7.5	379	4	US-09-491-577-4	Sequence 4, Appli
35	65	7.5	484	1	US-08-030-096-8	Sequence 8, Appli
36	65	7.5	1847	6	5256642-10	Patent No. 5256642
37	65	7.5	1847	6	5472939-10	Patent No. 5472939
38	65	7.5	2039	6	5256642-2	Patent No. 5256642
39	65	7.5	2039	6	5472939-2	Patent No. 5472939
40	65	7.5	2489	4	US-09-911-842A-5	Sequence 5, Appli
41	64.5	7.5	511	3	US-09-201-641-2	Sequence 2, Appli
42	64.5	7.5	511	4	US-09-323-998E-60	Sequence 60, Appl
43	64.5	7.5	859	4	US-09-134-000C-6022	Sequence 6022, Ap
44	64	7.4	483	3	US-09-108-020-16	Sequence 16, Appl
45	64	7.4	3635	4	US-09-845-583A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-994-962-2
; Sequence 2, Application US/08994962
; Patent No. 6284486
; GENERAL INFORMATION:
; APPLICANT: Olsen, Henrik S.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Oncogene Induced Secreted Protein I
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Ave., NW, Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/994,962
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/033,869
; FILING DATE: 20-DEC-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/037,388
; FILING DATE: 07-FEB-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0440002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 162 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-994-962-2

Query Match 100.0%; Score 862; DB 3; Length 162;
Best Local Similarity 100.0%; Pred. No. 8e-100;

```
  Matches  162;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MKGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSK 60

Qy         61 QLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 QLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQI 120

Qy        121 YYAGPVNNPEFTIPQGEYQVLLELYTEKRSTVACANATIMCS 162
              |||||||||||||||||||||||||||||||||||||||||
Db        121 YYAGPVNNPEFTIPQGEYQVLLELYTEKRSTVACANATIMCS 162


RESULT 2
US-08-994-962-3
; Sequence 3, Application US/08994962
; Patent No. 6284486
;  GENERAL INFORMATION:
;    APPLICANT:  Olsen, Henrik S.
;    APPLICANT:  Ruben, Steven M.
;    TITLE OF INVENTION:  Human Oncogene Induced Secreted Protein I
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Ave., NW, Suite 600
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/994,962
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/033,869
;      FILING DATE:  20-DEC-1996
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/037,388
;      FILING DATE:  07-FEB-1997
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K.
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1488.0440002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-2600
;      TELEFAX:  202-371-2540
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  133 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  protein
US-08-994-962-3

  Query Match             39.9%;  Score 344;  DB 3;  Length 133;
  Best Local Similarity   45.1%;  Pred. No. 4.2e-35;
  Matches   60;  Conservative   26;  Mismatches   47;  Indels    0;  Gaps    0;

Qy         27 WPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSKQLKSNINIRFGIILREDIKELFLDLA 86
              |||| || :  ||: |:|||| ||| ||:::||       :||  ::||: ||||:  :
Db          1 WPTHTVCKEENLEIYYKSCDPQQDFAFSIDRCSDVTTHTFDIRAAMVLRQSIKELYAKVD 60

Qy         87 LMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQIYYAGPVNNPEFTIPQGEYQVLLELYT 146
              |:  | :||::|  :|   | |  |||::||| :|| ||:       ||||:| :   |
Db         61 LIINGKTVLSYSETLCGPGLSKLIFCGKKKGEHLYYEGPITLGIKEIPQGDYTITARLTN 120

Qy        147 EKRSTVACANATI 159
              | |:|||||: |:
Db        121 EDRATVACADFTV 133


RESULT 3
US-08-994-962-17
; Sequence 17, Application US/08994962
; Patent No. 6284486
;  GENERAL INFORMATION:
;    APPLICANT:  Olsen, Henrik S.
;    APPLICANT:  Ruben, Steven M.
;    TITLE OF INVENTION:  Human Oncogene Induced Secreted Protein I
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Ave., NW, Suite 600
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/994,962
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/033,869
;      FILING DATE:  20-DEC-1996
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/037,388
;      FILING DATE:  07-FEB-1997
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K.
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1488.0440002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-2600
;      TELEFAX:  202-371-2540
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  60 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  peptide
US-08-994-962-17

  Query Match             10.9%;  Score 94;  DB 3;  Length 60;
  Best Local Similarity   33.3%;  Pred. No. 0.00027;
  Matches   43;  Conservative    3;  Mismatches   13;  Indels   70;  Gaps   11;

Qy         27 WPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSKQLKSNINIRFGIILREDIKELFLDLA 86
              |||||          |  |||| ||| ||   || :|:              ||||
Db          1 WPTHV--------CLEYSCDP-QDF-FS---CSIRLR--------------IKELL---- 29

Qy         87 LMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQIYYAGPVNNPEFTIPQGEYQVLLELYT 146
                                  |    |||  |||   | ||        ||||        |
Db         30 ----------------GVLSCLKFCG--KGE---YYGP-------IPQG--------YL 52

Qy        147 EKRSTVACA 155
              |:  |||||
Db         53 ER--TVACA 59
```

```
RESULT 4
US-08-969-644-16
; Sequence 16, Application US/08969644
; Patent No. 6096519
;  GENERAL INFORMATION:
;    APPLICANT:  Ratti, Giulio
;    APPLICANT:  Comanducci, Maurizio
;    APPLICANT:  Tecce, Mario F.
;    APPLICANT:  Giuliani, Marzia M.
;    TITLE OF INVENTION:  pCTD PLASMID ISOLATED FROM CHLAMYDIA
;    TITLE OF INVENTION:  TRACHOMATIS SEROTYPE D, ITS GENES AND PROTEINS ENCODED BY
;    TITLE OF INVENTION:  THEM; RECOMBINANT PLASMIDS FOR THE EXPRESSION OF SAID
;    TITLE OF INVENTION:  GENES IN HETEROLOGOUS SYSTEMS, PREPARATION OF SAID
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH & BIRCH
;      STREET:  301 N. Washington Street
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22046-0747
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/969,644
;      FILING DATE:  13-NOV-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/467,152
;      FILING DATE:
;      APPLICATION NUMBER:  US/07/661,820
;      FILING DATE:
;      APPLICATION NUMBER:  IT MI 91A000314
;      FILING DATE:  07-FEB-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Svensson, Leonard R.
;      REGISTRATION NUMBER:  30,330
;      REFERENCE/DOCKET NUMBER:  1267-202P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-241-1300
;      TELEFAX:  703-241-2848
;      TELEX:  248345
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  264 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-969-644-16

  Query Match             8.5%;  Score 73.5;  DB 3;  Length 264;
  Best Local Similarity   27.6%;  Pred. No. 0.83;
  Matches   35;  Conservative   13;  Mismatches   50;  Indels   29;  Gaps    5;

Qy         10 LWTLIFPSCSGGGGGKAWPTHVVCSDS---GLEVLYQSCDPLQDFGFSVEKCSKQLKSNI 66
                | ||:| |  || |     :| |: :  | :||   ||                :||:
Db          1 LHTLVFCSFKGGTGKTTLSLNVGCNLAQFLGKKVLLADLDP--------------QSNL 45

Qy         67 NIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAA-----LPKFSFCGRRKGEQIY 121
              :   | :| | | |  |:   |     :     |||      |   ||   : |
Db         46 SSGLGASVRSDQKGLH-DIVYTSNDLKSI-----ICETKKDSVDLIPASFSSEQFRELDI 99

Qy        122 YAGPVNN 128
                : || ||
Db        100 HRGPSNN 106


RESULT 5
US-08-444-189-16
; Sequence 16, Application US/08444189
; Patent No. 6110705
;  GENERAL INFORMATION:
;    APPLICANT:  Ratti, Giulio
;    APPLICANT:  Comanducci, Maurizio
;    APPLICANT:  Tecce, Mario F.
;    APPLICANT:  Giuliani, Marzia M.
;    TITLE OF INVENTION:  pCTD PLASMID ISOLATED FROM CHLAMYDIA
;    TITLE OF INVENTION:  TRACHOMATIS SEROTYPE D, ITS GENES AND PROTEINS ENCODED BY
;    TITLE OF INVENTION:  THEM; RECOMBINANT PLASMIDS FOR THE EXPRESSION OF SAID
;    TITLE OF INVENTION:  GENES IN HETEROLOGOUS SYSTEMS, PREPARATION OF SAID
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH & BIRCH
;      STREET:  301 N. Washington Street
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22046-0747
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/444,189
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/180,528
;      FILING DATE:
;      APPLICATION NUMBER:  US/07/991,512
;      FILING DATE:
;      APPLICATION NUMBER:  US/07/661,820
;      FILING DATE:
;      APPLICATION NUMBER:  IT MI 91A000314
;      FILING DATE:  07-FEB-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Svensson, Leonard R.
;      REGISTRATION NUMBER:  30,330
;      REFERENCE/DOCKET NUMBER:  1267-202P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-241-1300
;      TELEFAX:  703-241-2848
;      TELEX:  248345
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  264 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-444-189-16

  Query Match             8.5%;  Score 73.5;  DB 3;  Length 264;
  Best Local Similarity   27.6%;  Pred. No. 0.83;
  Matches   35;  Conservative   13;  Mismatches   50;  Indels   29;  Gaps    5;

Qy         10 LWTLIFPSCSGGGGGKAWPTHVVCSDS---GLEVLYQSCDPLQDFGFSVEKCSKQLKSNI 66
                | ||:| |  || |     :| |: :  | :||   ||                :||:
Db          1 LHTLVFCSFKGGTGKTTLSLNVGCNLAQFLGKKVLLADLDP--------------QSNL 45

Qy         67 NIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAA-----LPKFSFCGRRKGEQIY 121
              :   | :| | | |  |:   |     :     |||      |   ||   : |
Db         46 SSGLGASVRSDQKGLH-DIVYTSNDLKSI-----ICETKKDSVDLIPASFSSEQFRELDI 99

Qy        122 YAGPVNN 128
                : || ||
Db        100 HRGPSNN 106


RESULT 6
US-08-468-544-16
; Sequence 16, Application US/08468544
```

```
; Patent No. 6248563
;  GENERAL INFORMATION:
;    APPLICANT:  Ratti, Giulio
;    APPLICANT:  Comanducci, Maurizio
;    APPLICANT:  Tecce, Mario F.
;    APPLICANT:  Giuliani, Marzia M.
;    TITLE OF INVENTION:  pCTD PLASMID ISOLATED FROM CHLAMYDIA
;    TITLE OF INVENTION:  TRACHOMATIS SEROTYPE D, ITS GENES AND PROTEINS ENCODED BY
;    TITLE OF INVENTION:  THEM; RECOMBINANT PLASMIDS FOR THE EXPRESSION OF SAID
;    TITLE OF INVENTION:  GENES IN HETEROLOGOUS SYSTEMS, PREPARATION OF SAID
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH & BIRCH
;      STREET:  301 N. Washington Street
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22046-0747
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/468,544
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/661,820
;      FILING DATE:  28-FEB-1991
;      APPLICATION NUMBER:  IT MI 91A000314
;      FILING DATE:  07-FEB-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Svensson, Leonard R.
;      REGISTRATION NUMBER:  30,330
;      REFERENCE/DOCKET NUMBER:  1267-202P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-241-1300
;      TELEFAX:  703-241-2848
;      TELEX:  248345
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  264 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-468-544-16

  Query Match             8.5%;  Score 73.5;  DB 3;  Length 264;
  Best Local Similarity   27.6%;  Pred. No. 0.83;
  Matches   35;  Conservative   13;  Mismatches   50;  Indels   29;  Gaps    5;

Qy         10 LWTLIFPSCSGGGGGKAWPTHVVCSDS---GLEVLYQSCDPLQDFGFSVEKCSKQLKSNI 66
                | ||:| |  || |      :| |: :  | :||    ||               :||:
Db          1 LHTLVFCSFKGGTGKTTLSLNVGCNLAQFLGKKVLLADLDP--------------QSNL 45

Qy         67 NIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAA-----LPKFSFCGRRKGEQIY 121
              :   |  :| | | |  |:   |      :     |||       |   ||   :  |
Db         46 SSGLGASVRSDQKGLH-DIVYTSNDLKSI-----ICETKKDSVDLIPASFSSEQFRELDI 99

Qy        122 YAGPVNN 128
              : || ||
Db        100 HRGPSNN 106


RESULT 7
US-08-220-151-22
; Sequence 22, Application US/08220151
; Patent No. 5529780
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    APPLICANT:  Limbach, Keith J.
;    TITLE OF INVENTION:  NUCLEOTIDE AND AMINO ACID SEQUENCES OF
;    TITLE OF INVENTION:  CANINE HERPESVIRUS gB, gC AND gD AND USES THEREFOR
;    NUMBER OF SEQUENCES:  91
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/220,151
;      FILING DATE:  30-MAR-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Frommer, William S.
;      REGISTRATION NUMBER:  25,506
;      REFERENCE/DOCKET NUMBER:  454310-2540
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;      TELEX:  425066 CURTMS
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  442 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FRAGMENT TYPE:  N-terminal
US-08-220-151-22

  Query Match             8.5%;  Score 73.5;  DB 1;  Length 442;
  Best Local Similarity   25.4%;  Pred. No. 1.7;
  Matches   43;  Conservative   17;  Mismatches   62;  Indels   47;  Gaps    8;

Qy          2 KGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSKQ 61
                | ::| |  |  | |:|       | | |        :| |  |:|   |    |
Db        161 KTYSARL-TWFKIMPTC-------ATPIH--------DVSYMKCNPKLSFAMCDE----- 199

Qy         62 LKSNINIRFGII--LREDIKELFLDLALMSQGSSVL-------------NFSYPICEAA 105
                :|:|  :  :|    |   || | ||  :  :| |             :|| |
Db        200 -RSDILWQASLITMAAETDDELGLVLAAPAHSASGLYRRVIEIDGRRIYTDFSVTIPSER 258

Qy        106 LP-----KFSFCGRRKGEQIYYAGPVNN----PEFTIPQGEYQVLLELY 145
                |     |   | |  : |  | |       ||  ||||:   :: :
Db        259 CPIAFELNFGNPDRCKTPEQYSRGEVFTRRFLGEFNFPQGEHMTWVKFW 307


RESULT 8
US-08-413-118-22
; Sequence 22, Application US/08413118
; Patent No. 5688920
;  GENERAL INFORMATION:
;    APPLICANT:  PAOLETTI, ENZO
;    APPLICANT:  LIMBACH, KEITH J.
;    TITLE OF INVENTION:  NUCLEOTIDE AND AMINO ACID SEQUENCES OF
;    TITLE OF INVENTION:  CANINE HERPESVIRUS gB, gC, AND gD AND USES THEREFOR
;    NUMBER OF SEQUENCES:  128
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CURTIS, MORRIS & SAFFORD, P.C.
;      STREET:  530 FIFTH AVENUE, 25TH FLOOR
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/413,118
;       FILING DATE:  29-MAR-1995
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/220,151
;       FILING DATE:  30-MAR-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  FROMMER, WILLIAM S.
;       REGISTRATION NUMBER:  25,506
;       REFERENCE/DOCKET NUMBER:  454310-2670
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 840-3333
;       TELEFAX:  (212) 840-0712
;   INFORMATION FOR SEQ ID NO:  22:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  442 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     FRAGMENT TYPE:  N-terminal
US-08-413-118-22

  Query Match             8.5%;  Score 73.5;  DB 1;  Length 442;
  Best Local Similarity   25.4%;  Pred. No. 1.7;
  Matches   43;  Conservative   17;  Mismatches   62;  Indels   47;  Gaps    8;

Qy          2 KGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSKQ 61
              | ::| |  |  | |:|       | | |        :| |  |:|   |     |
Db        161 KTYSARL-TWFKIMPTC-------ATPIH--------DVSYMKCNPKLSFAMCDE----- 199

Qy         62 LKSNINIRFGII--LREDIKELFLDLALMSQGSSVL-------------NFSYPICEAA 105
               :|:|  :  :|    |   || | ||  :  :| |                :||  |
Db        200 -RSDILWQASLITMAAETDDELGLVLAAPAHSASGLYRRVIEIDGRRIYTDFSVTIPSER 258

Qy        106 LP-----KFSFCGRRKGEQIYYAGPVNN----PEFTIPQGEYQVLLELY 145
              |       |    | |  : |  | |        ||  ||||:   :: :
Db        259 CPIAFELNFGNPDRCKTPEQYSRGEVFTRRFLGEFNFPQGEHMTWVKFW 307


RESULT 9
US-08-473-446-22
; Sequence 22, Application US/08473446
; Patent No. 6017542
;  GENERAL INFORMATION:
;    APPLICANT:  PAOLETTI, ENZO
;    APPLICANT:  LIMBACH, KEITH J.
;    TITLE OF INVENTION:  NUCLEOTIDE AND AMINO ACID SEQUENCES OF
;    TITLE OF INVENTION:  CANINE HERPESVIRUS gB, gC, AND gD AND USES THEREFOR
;    NUMBER OF SEQUENCES:  128
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CURTIS, MORRIS & SAFFORD, P.C.
;      STREET:  530 FIFTH AVENUE, 25TH FLOOR
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/473,446
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/413,118
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FROMMER, WILLIAM S.
;      REGISTRATION NUMBER:  25,506
;      REFERENCE/DOCKET NUMBER:  454310-2670
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  442 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FRAGMENT TYPE:  N-terminal
US-08-473-446-22

  Query Match             8.5%;  Score 73.5;  DB 3;  Length 442;
  Best Local Similarity   25.4%;  Pred. No. 1.7;
  Matches   43;  Conservative   17;  Mismatches   62;  Indels   47;  Gaps    8;

Qy          2 KGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSKQ 61
              | ::| |  |  | |:|       | | |        :| |  |:|   |     |
Db        161 KTYSARL-TWFKIMPTC-------ATPIH--------DVSYMKCNPKLSFAMCDE----- 199

Qy         62 LKSNINIRFGII--LREDIKELFLDLALMSQGSSVL-------------NFSYPICEAA 105
               :|:|  :  :|    |   || | ||  :  :| |                :||  |
Db        200 -RSDILWQASLITMAAETDDELGLVLAAPAHSASGLYRRVIEIDGRRIYTDFSVTIPSER 258

Qy        106 LP-----KFSFCGRRKGEQIYYAGPVNN----PEFTIPQGEYQVLLELY 145
              |       |    | |  : |  | |        ||  ||||:   :: :
Db        259 CPIAFELNFGNPDRCKTPEQYSRGEVFTRRFLGEFNFPQGEHMTWVKFW 307


RESULT 10
5470718-3
;Patent No. 5470718
;    APPLICANT: O'CALLAGHAN, DENNIS J.
;    TITLE OF INVENTION: EQUINE HERPESVIRUS TYPE 1 GLYCOPROTEIN
;D NUCLEIC ACIDS
;    NUMBER OF SEQUENCES: 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/954,417
;      FILING DATE: 30-SEP-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 561,553
;      FILING DATE: 01-AUG-1990
;SEQ ID NO:3:
;      LENGTH: 442
5470718-3

  Query Match             8.5%;  Score 73.5;  DB 6;  Length 442;
  Best Local Similarity   25.4%;  Pred. No. 1.7;
  Matches   43;  Conservative   17;  Mismatches   62;  Indels   47;  Gaps    8;

Qy          2 KGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSKQ 61
              | ::| |  |  | |:|       | | |        :| |  |:|   |     |
Db        161 KTYSARL-TWFKIMPTC-------ATPIH--------DVSYMKCNPKLSFAMCDE----- 199

Qy         62 LKSNINIRFGII--LREDIKELFLDLALMSQGSSVL-------------NFSYPICEAA 105
               :|:|  :  :|    |   || | ||  :  :| |                :||  |
Db        200 -RSDILWQASLITMAAETDDELGLVLAAPAHSASGLYRRVIEIDGRRIYTDFSVTIPSER 258

Qy        106 LP-----KFSFCGRRKGEQIYYAGPVNN----PEFTIPQGEYQVLLELY 145
              |       |    | |  : |  | |        ||  ||||:   :: :
Db        259 CPIAFELNFGNPDRCKTPEQYSRGEVFTRRFLGEFNFPQGEHMTWVKFW 307


RESULT 11
US-09-252-991A-26468
; Sequence 26468, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
```

```
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 26468
;   LENGTH: 787
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26468

  Query Match             8.5%;  Score 73;  DB 4;  Length 787;
  Best Local Similarity   22.3%;  Pred. No. 4.6;
  Matches   27;  Conservative   14;  Mismatches   46;  Indels   34;  Gaps    3;

Qy         10 LWTLIFPSCSGGGGGKAWPTHVVC--------------SDSGLEVL-----------YQS 44
              ||||| |     ||  :|     :|              :| | : |
Db        273 LWTLIRPGKPDGGNRRATILKSMCRIPTIGPVRAERLLNDFGEDFLATMLVDNVSEFINL 332

Qy         45 CDPLQDFGFSVEKCSKQLKSNINIRFGI---------ILREDIKELFLDLALMSQGSSVL 95
                |  || ||  :  :  :|  || ||           ::  :   : || :: :|
Db        333 MDAKGDFVFSDRQAKRMERSMANIEFGFGEGGYQPTEFIKRQLPNGYFDLLVVDEGHEYK 392

Qy         96 N 96
              |
Db        393 N 393


RESULT 12
5470718-2
;Patent No. 5470718
;    APPLICANT: O'CALLAGHAN, DENNIS J.
;    TITLE OF INVENTION: EQUINE HERPESVIRUS TYPE 1 GLYCOPROTEIN
;D NUCLEIC ACIDS
;    NUMBER OF SEQUENCES: 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/954,417
;      FILING DATE: 30-SEP-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 561,553
;      FILING DATE: 01-AUG-1990
;SEQ ID NO:2:
;      LENGTH: 442
5470718-2

  Query Match             8.3%;  Score 71.5;  DB 6;  Length 442;
  Best Local Similarity   25.4%;  Pred. No. 3.1;
  Matches   43;  Conservative   16;  Mismatches   63;  Indels   47;  Gaps    8;

Qy          2 KGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSKQ 61
              | : | |  |  | | |:|       | | |         :| |  |:|   |    |
Db        161 KTYPARL-TWFKIMPTC-------ATPIH--------DVSYMKCNPKLSFAMCDE----- 199

Qy         62 LKSNINIRFGII--LREDIKELFLDLALMSQGSSVL-------------NFSYPICEAA 105
                :|:|  :  :|    |   || | ||  :  :| |              :||  |
Db        200 -RSDILWQASLITMAAETDDELGLVLAAPAHSASGLYRRVIEIDGRRIYTDFSVTIPSER 258

Qy        106 LP-----KFSFCGRRKGEQIYYAGPVNN----PEFTIPQGEYQVLLELY 145
              |     |   | | : | | |       ||  ||||:   :: :
Db        259 CPIAFELNFGNPDRCKTPEQYSRGEVFTRRFLGEFNFPQGEHMTWVKFW 307


RESULT 13
US-08-530-950-13
; Sequence 13, Application US/08530950
; Patent No. 5736381
;  GENERAL INFORMATION:
;    APPLICANT:  Davis, Roger J.
;    APPLICANT:  Raingeaud, Joel
;    APPLICANT:  Gupta, Shashi
;    APPLICANT:  Derijard, Benoit
;    TITLE OF INVENTION:  CYTOKINE-, STRESS-, AND
;    TITLE OF INVENTION:  ONCOPROTEIN-ACTIVATED HUMAN PROTEIN KINASE
;    TITLE OF INVENTION:  KINASES
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/530,950
;      FILING DATE:  19-SEP-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fasse, J. Peter
;      REGISTRATION NUMBER:  32,983
;      REFERENCE/DOCKET NUMBER:  07917/010001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  668 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
US-08-530-950-13

  Query Match             8.2%;  Score 71;  DB 1;  Length 668;
  Best Local Similarity   24.3%;  Pred. No. 6.4;
  Matches   33;  Conservative   24;  Mismatches   61;  Indels   18;  Gaps    5;

Qy         28 PTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSKQLKSNINIRFGIILREDIKELFLDLAL 87
              ||:::|| :       |    | ||| |    :   |:||  :   :  | || |  | |
Db        488 PTNILCSAN------QGTVKLCDFGVSGNLVASLAKTNIGCQ-SYMAPERIKSLNPDRAT 540

Qy         88 MSQGSSVLNFSYPICEAALPKFSFCGRRKGEQIY--YAGPVNNPEFTIPQGEY------- 138
                :  | : :    | | || :: :      : |:   :  |: |   :|  ::
Db        541 YTVQSDIWSLGLSILEMALGRYPY-PPETYDNIFSQLSAIVDGPPPRLPSDKFSSDAQDF 599

Qy        139 -QVLLELYTEKRSTVA 153
               : |:   |:| | |
Db        600 VSLCLQKIPERRPTYA 615


RESULT 14
US-09-149-879-13
; Sequence 13, Application US/09149879
; Patent No. 6174676
;  GENERAL INFORMATION:
;    APPLICANT:  Davis, Roger J.
;    APPLICANT:  Raingeaud, Joel
;    APPLICANT:  Gupta, Shashi
;    APPLICANT:  Derijard, Benoit
;    TITLE OF INVENTION:  CYTOKINE-, STRESS-, AND
;    TITLE OF INVENTION:  ONCOPROTEIN-ACTIVATED HUMAN PROTEIN KINASE
;    TITLE OF INVENTION:  KINASES
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
```

```
;       STATE:  MA
;       COUNTRY:  USA
;       ZIP:  02110-2804
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/149,879
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/530,950
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Fasse, J. Peter
;       REGISTRATION NUMBER:  32,983
;       REFERENCE/DOCKET NUMBER:  07917/010001
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  617/542-5070
;       TELEFAX:  617/542-8906
;       TELEX:  200154
;   INFORMATION FOR SEQ ID NO:  13:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  668 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  not relevant
;       TOPOLOGY:  linear
US-09-149-879-13

  Query Match             8.2%;  Score 71;  DB 3;  Length 668;
  Best Local Similarity   24.3%;  Pred. No. 6.4;
  Matches   33;  Conservative   24;  Mismatches   61;  Indels   18;  Gaps    5;

Qy         28 PTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSKQLKSNINIRFGIILREDIKELFLDLAL 87
              ||:::|| :       |    | ||| |    :   |:||  :   :  | || |  | |
Db        488 PTNILCSAN------QGTVKLCDFGVSGNLVASLAKTNIGCQ-SYMAPERIKSLNPDRAT 540

Qy         88 MSQGSSVLNFSYPICEAALPKFSFCGRRKGEQIY--YAGPVNNPEFTIPQGEY------- 138
                :  | : :    | | || :: :      : |:   :  |: |   :|  ::
Db        541 YTVQSDIWSLGLSILEMALGRYPY-PPETYDNIFSQLSAIVDGPPPRLPSDKFSSDAQDF 599

Qy        139 -QVLLELYTEKRSTVA 153
               : |:   |:| | |
Db        600 VSLCLQKIPERRPTYA 615


RESULT 15
US-09-057-009-13
; Sequence 13, Application US/09057009
; Patent No. 6541605
;  GENERAL INFORMATION:
;    APPLICANT:  Davis, Roger J.
;    APPLICANT:  Raingeaud, Joel
;    APPLICANT:  Gupta, Shashi
;    APPLICANT:  Derijard, Benoit
;    TITLE OF INVENTION:  CYTOKINE-, STRESS-, AND
;    TITLE OF INVENTION:  ONCOPROTEIN-ACTIVATED HUMAN PROTEIN KINASE
;    TITLE OF INVENTION:  KINASES
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/057,009
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/530,950
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fasse, J. Peter
;      REGISTRATION NUMBER:  32,983
;      REFERENCE/DOCKET NUMBER:  07917/010001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  668 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  No. 6541605 Relevant
;      TOPOLOGY:  linear
US-09-057-009-13

  Query Match             8.2%;  Score 71;  DB 4;  Length 668;
  Best Local Similarity   24.3%;  Pred. No. 6.4;
  Matches   33;  Conservative   24;  Mismatches   61;  Indels   18;  Gaps    5;

Qy         28 PTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSKQLKSNINIRFGIILREDIKELFLDLAL 87
              ||:::|| :       |    | ||| |    :   |:||  :   :  | || |  | |
Db        488 PTNILCSAN------QGTVKLCDFGVSGNLVASLAKTNIGCQ-SYMAPERIKSLNPDRAT 540

Qy         88 MSQGSSVLNFSYPICEAALPKFSFCGRRKGEQIY--YAGPVNNPEFTIPQGEY------- 138
                :  | : :    | | || :: :      : |:   :  |: |   :|  ::
Db        541 YTVQSDIWSLGLSILEMALGRYPY-PPETYDNIFSQLSAIVDGPPPRLPSDKFSSDAQDF 599

Qy        139 -QVLLELYTEKRSTVA 153
               : |:   |:| | |
Db        600 VSLCLQKIPERRPTYA 615


Search completed:  April 29, 2004, 15:23:47
Job time :  23 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: April 29, 2004, 15:23:51 ; Search time 42 Seconds
(without alignments)
1069.160 Million cell updates/sec

Title: US-09-899-917-2
Perfect score: 862
Sequence: 1 MKGFTATLFLWTLIFPSCSG.....ELYTEKRSTVACANATIMCS 162

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1138120 seqs, 277189581 residues

Total number of hits satisfying chosen parameters: 1138120

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	862	100.0	162	9	US-09-899-917-2	Sequence 2, Appli
2	862	100.0	162	12	US-10-000-066-13	Sequence 13, Appl
3	850	98.6	162	10	US-09-372-348-19	Sequence 19, Appl
4	613	71.1	162	10	US-09-372-348-27	Sequence 27, Appl
5	613	71.1	162	12	US-10-000-066-10	Sequence 10, Appl
6	344	39.9	133	9	US-09-899-917-3	Sequence 3, Appli
7	336	39.0	160	10	US-09-372-348-29	Sequence 29, Appl
8	106	12.3	160	12	US-10-000-066-1	Sequence 1, Appli
9	101.5	11.8	160	10	US-09-372-348-23	Sequence 23, Appl
10	101.5	11.8	160	12	US-10-000-066-7	Sequence 7, Appli
11	96.5	11.2	141	10	US-09-372-348-26	Sequence 26, Appl
12	94	10.9	60	9	US-09-899-917-17	Sequence 17, Appl
13	81.5	9.5	555	12	US-10-424-599-214951	Sequence 214951,
14	81	9.4	170	12	US-10-424-599-199134	Sequence 199134,
15	78	9.0	876	12	US-10-282-122A-77326	Sequence 77326, A
16	76.5	8.9	1177	12	US-10-282-122A-45910	Sequence 45910, A
17	75.5	8.8	342	12	US-10-282-122A-54490	Sequence 54490, A
18	75.5	8.8	1241	12	US-10-282-122A-51957	Sequence 51957, A
19	75	8.7	505	14	US-10-205-823-427	Sequence 427, App
20	75	8.7	505	14	US-10-354-358-98	Sequence 98, Appl
21	75	8.7	522	12	US-10-087-192-84	Sequence 84, Appl
22	75	8.7	524	9	US-09-925-300-1538	Sequence 1538, Ap
23	72.5	8.4	1172	12	US-10-363-616-446	Sequence 446, App
24	72.5	8.4	1178	12	US-10-363-616-445	Sequence 445, App
25	72	8.4	266	12	US-10-425-114-53807	Sequence 53807, A
26	72	8.4	266	12	US-10-425-114-53809	Sequence 53809, A
27	72	8.4	266	12	US-10-425-114-68478	Sequence 68478, A
28	71.5	8.3	313	15	US-10-041-615-28	Sequence 28, Appl
29	71	8.2	668	9	US-09-761-569-13	Sequence 13, Appl
30	71	8.2	668	9	US-09-801-368-238	Sequence 238, App
31	71	8.2	668	14	US-10-081-119-20	Sequence 20, Appl
32	71	8.2	668	15	US-10-369-493-22148	Sequence 22148, A
33	71	8.2	1115	14	US-10-260-715-6	Sequence 6, Appli
34	71	8.2	1151	14	US-10-260-715-4	Sequence 4, Appli
35	70.5	8.2	114	10	US-09-372-348-25	Sequence 25, Appl
36	70.5	8.2	114	12	US-10-000-066-4	Sequence 4, Appli
37	70.5	8.2	218	12	US-10-335-977-6276	Sequence 6276, Ap
38	70.5	8.2	495	12	US-10-424-599-234625	Sequence 234625,
39	70	8.1	215	12	US-10-282-122A-77503	Sequence 77503, A
40	70	8.1	291	15	US-10-264-237-1641	Sequence 1641, Ap
41	70	8.1	462	11	US-09-833-245-1692	Sequence 1692, Ap
42	70	8.1	515	15	US-10-369-493-17475	Sequence 17475, A
43	70	8.1	950	15	US-10-369-493-21918	Sequence 21918, A
44	69.5	8.1	191	12	US-10-424-599-234053	Sequence 234053,
45	69.5	8.1	373	15	US-10-434-418-2	Sequence 2, Appli

ALIGNMENTS

Instant appl.

```
RESULT 1
US-09-899-917-2
; Sequence 2, Application US/09899917
; Patent No. US20020119552A1
;  GENERAL INFORMATION:
;       APPLICANT: Olsen, Henrik S.
;                  Ruben, Steven M.
;       TITLE OF INVENTION: Human Oncogene Induced Secreted Protein I
;       NUMBER OF SEQUENCES: 17
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Sterne, Kessler Goldstein & Fox P.L.L.C.
;            STREET: 1100 New York Ave., NW, Suite 600
;            CITY: Washington
;            STATE: DC
;            COUNTRY: USA
;            ZIP: 20005-3934
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/899,917
;            FILING DATE: 09-Jul-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/994,962
;            FILING DATE: <Unknown>
;            APPLICATION NUMBER: US 60/037,388
;            FILING DATE: 07-FEB-1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Steffe, Eric K.
;            REGISTRATION NUMBER: 36,688
;            REFERENCE/DOCKET NUMBER: 1488.0440002
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 202-371-2600
;            TELEFAX: 202-371-2540
;  INFORMATION FOR SEQ ID NO: 2:
```

```
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 162 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-899-917-2

  Query Match             100.0%;  Score 862;  DB 9;  Length 162;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-93;
  Matches  162;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MKGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSK 60

Qy         61 QLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 QLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQI 120

Qy        121 YYAGPVNNPEFTIPQGEYQVLLELYTEKRSTVACANATIMCS 162
              |||||||||||||||||||||||||||||||||||||||||
Db        121 YYAGPVNNPEFTIPQGEYQVLLELYTEKRSTVACANATIMCS 162


RESULT 2
US-10-000-066-13
; Sequence 13, Application US/10000066
; Publication No. US20020086364A1
; GENERAL INFORMATION:
;  APPLICANT: HONJO, TASUKU
;  APPLICANT:  KATO, KEIZO
;  APPLICANT:  TADA, HIDEAKI
;  TITLE OF INVENTION: POLYPEPTIDE, cDNA ENCODING THE SAME, AND USE OF THEM
;  FILE REFERENCE: Q58771
;  CURRENT APPLICATION NUMBER: US/10/000,066
;  CURRENT FILING DATE:  2001-12-04
;  PRIOR APPLICATION NUMBER: 09/529,064
;  PRIOR FILING DATE: 2000-04-17
;  PRIOR APPLICATION NUMBER: PCT/JP98/04515
;  PRIOR FILING DATE: 1998-10-06
;  PRIOR APPLICATION NUMBER: HEI-9-274673
;  PRIOR FILING DATE: 1997-10-07
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
;   LENGTH: 162
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-000-066-13

  Query Match             100.0%;  Score 862;  DB 12;  Length 162;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-93;
  Matches  162;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MKGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSK 60

Qy         61 QLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 QLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQI 120

Qy        121 YYAGPVNNPEFTIPQGEYQVLLELYTEKRSTVACANATIMCS 162
              |||||||||||||||||||||||||||||||||||||||||
Db        121 YYAGPVNNPEFTIPQGEYQVLLELYTEKRSTVACANATIMCS 162


RESULT 3
US-09-372-348-19
; Sequence 19, Application US/09372348A
; Publication No. US20030028005A1
; GENERAL INFORMATION:
;  APPLICANT: Hedrick, Joseph A.
;  APPLICANT: Wang, Luquan
;  APPLICANT: Zlotnik, Albert
;  APPLICANT: Murgolo, Nicholas J.
;  APPLICANT: Greene, Jonathan R.
;  APPLICANT: Johnston, James A.
;  APPLICANT: Bazan, J. Fernando
;  TITLE OF INVENTION: Mammalian Proteins; Related Reagents and Methods
;  FILE REFERENCE: DX0884K(2d)
;  CURRENT APPLICATION NUMBER: US/09/372,348A
;  CURRENT FILING DATE: 1999-08-11
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
;   LENGTH: 162
;   TYPE: PRT
;   ORGANISM: primate
US-09-372-348-19

  Query Match             98.6%;  Score 850;  DB 10;  Length 162;
  Best Local Similarity   98.8%;  Pred. No. 4.8e-92;
  Matches  160;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MKGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSK 60
              |||||||||||| |||||||||||||||||||| |||||||||||||||||||||||||||
Db          1 MKGFTATLFLWTXIFPSCSGGGGGKAWPTHVVCSDSXLEVLYQSCDPLQDFGFSVEKCSK 60

Qy         61 QLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 QLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQI 120

Qy        121 YYAGPVNNPEFTIPQGEYQVLLELYTEKRSTVACANATIMCS 162
              |||||||||||||||||||||||||||||||||||||||||
Db        121 YYAGPVNNPEFTIPQGEYQVLLELYTEKRSTVACANATIMCS 162


RESULT 4
US-09-372-348-27
; Sequence 27, Application US/09372348A
; Publication No. US20030028005A1
; GENERAL INFORMATION:
;  APPLICANT: Hedrick, Joseph A.
;  APPLICANT: Wang, Luquan
;  APPLICANT: Zlotnik, Albert
;  APPLICANT: Murgolo, Nicholas J.
;  APPLICANT: Greene, Jonathan R.
;  APPLICANT: Johnston, James A.
;  APPLICANT: Bazan, J. Fernando
;  TITLE OF INVENTION: Mammalian Proteins; Related Reagents and Methods
;  FILE REFERENCE: DX0884K(2d)
;  CURRENT APPLICATION NUMBER: US/09/372,348A
;  CURRENT FILING DATE: 1999-08-11
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
;   LENGTH: 162
;   TYPE: PRT
;   ORGANISM: rodent
US-09-372-348-27

  Query Match             71.1%;  Score 613;  DB 10;  Length 162;
  Best Local Similarity   66.7%;  Pred. No. 5e-64;
  Matches  108;  Conservative   22;  Mismatches   32;  Indels    0;  Gaps    0;

Qy          1 MKGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSK 60
              | |  | | :| |  || |  |    || |  |:  ||||:|||||||||| |:::|||
Db          1 MNGVAAALLVWILTSPSSSDHGSENGWPKHTACNSGGLEVVYQSCDPLQDFGLSIDQCSK 60

Qy         61 QLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQI 120
              |::||:|||||||||:||::||||: ||::|||:||:|||:||   |||||||||||||||
Db         61 QIQSNLNIRFGIILRQDIRKLFLDITLMAKGSSILNYSYPLCEEDQPKFSFCGRRKGEQI 120

Qy        121 YYAGPVNNPEFTIPQGEYQVLLELYTEKRSTVACANATIMCS 162
```

```
                ||||||||||   :||||||:|||||  | |:||||||||:  |
Db        121 YYAGPVNNPGLDVPQGEYQLLLELYNENRATVACANATVTSS 162


RESULT 5
US-10-000-066-10
; Sequence 10, Application US/10000066
; Publication No. US20020086364A1
; GENERAL INFORMATION:
;  APPLICANT: HONJO, TASUKU
;  APPLICANT:  KATO, KEIZO
;  APPLICANT:  TADA, HIDEAKI
;  TITLE OF INVENTION: POLYPEPTIDE, cDNA ENCODING THE SAME, AND USE OF THEM
;  FILE REFERENCE: Q58771
;  CURRENT APPLICATION NUMBER: US/10/000,066
;  CURRENT FILING DATE:  2001-12-04
;  PRIOR APPLICATION NUMBER: 09/529,064
;  PRIOR FILING DATE: 2000-04-17
;  PRIOR APPLICATION NUMBER: PCT/JP98/04515
;  PRIOR FILING DATE: 1998-10-06
;  PRIOR APPLICATION NUMBER: HEI-9-274673
;  PRIOR FILING DATE: 1997-10-07
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
;   LENGTH: 162
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-000-066-10

  Query Match             71.1%;  Score 613;  DB 12;  Length 162;
  Best Local Similarity   66.7%;  Pred. No. 5e-64;
  Matches  108;  Conservative   22;  Mismatches   32;  Indels    0;  Gaps    0;

Qy          1 MKGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSK 60
              |  |  | | :| |  || |  |    || |  |:  ||||:||||||||||| |:::|||
Db          1 MNGVAAALLVWILTSPSSSDHGSENGWPKHTACNSGGLEVVYQSCDPLQDFGLSIDQCSK 60

Qy         61 QLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQI 120
              |::||:|||||||||:||::|||| ||::|||:||:|||:||   ||||||||||||||
Db         61 QIQSNLNIRFGIILRQDIRKLFLDITLMAKGSSILNYSYPLCEEDQPKFSFCGRRKGEQI 120

Qy        121 YYAGPVNNPEFTIPQGEYQVLLELYTEKRSTVACANATIMCS 162
              |||||||||   :||||||:|||||  | |:||||||||:  |
Db        121 YYAGPVNNPGLDVPQGEYQLLLELYNENRATVACANATVTSS 162


RESULT 6
US-09-899-917-3
; Sequence 3, Application US/09899917
; Patent No. US20020119552A1
;   GENERAL INFORMATION:
;        APPLICANT: Olsen, Henrik S.
;                   Ruben, Steven M.
;        TITLE OF INVENTION: Human Oncogene Induced Secreted Protein I
;        NUMBER OF SEQUENCES: 17
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Sterne, Kessler Goldstein & Fox P.L.L.C.
;             STREET: 1100 New York Ave., NW, Suite 600
;             CITY: Washington
;             STATE: DC
;             COUNTRY: USA
;             ZIP: 20005-3934
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/899,917
;             FILING DATE: 09-Jul-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/994,962
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: US 60/037,388
;             FILING DATE: 07-FEB-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Steffe, Eric K.
;             REGISTRATION NUMBER: 36,688
;             REFERENCE/DOCKET NUMBER: 1488.0440002
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 202-371-2600
;             TELEFAX: 202-371-2540
;   INFORMATION FOR SEQ ID NO: 3:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 133 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: No. US20020119552A1 Relevant
;             TOPOLOGY: No. US20020119552A1 Relevant
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-899-917-3

  Query Match             39.9%;  Score 344;  DB 9;  Length 133;
  Best Local Similarity   45.1%;  Pred. No. 2.4e-32;
  Matches   60;  Conservative   26;  Mismatches   47;  Indels    0;  Gaps    0;

Qy         27 WPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSKQLKSNINIRFGIILREDIKELFLDLA 86
              |||| || :  ||: |:|||| ||| ||:::||       :||  ::||: ||||:  :
Db          1 WPTHTVCKEENLEIYYKSCDPQQDFAFSIDRCSDVTTHTFDIRAAMVLRQSIKELYAKVD 60

Qy         87 LMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQIYYAGPVNNPEFTIPQGEYQVLLELYT 146
              |:  | :||::|  :|   | |  |||::||| :|| ||:       ||||:| :   |
Db         61 LIINGKTVLSYSETLCGPGLSKLIFCGKKKGEHLYYEGPITLGIKEIPQGDYTITARLTN 120

Qy        147 EKRSTVACANATI 159
              | |:|||||: |:
Db        121 EDRATVACADFTV 133


RESULT 7
US-09-372-348-29
; Sequence 29, Application US/09372348A
; Publication No. US20030028005A1
; GENERAL INFORMATION:
;  APPLICANT: Hedrick, Joseph A.
;  APPLICANT: Wang, Luquan
;  APPLICANT: Zlotnik, Albert
;  APPLICANT: Murgolo, Nicholas J.
;  APPLICANT: Greene, Jonathan R.
;  APPLICANT: Johnston, James A.
;  APPLICANT: Bazan, J. Fernando
;  TITLE OF INVENTION: Mammalian Proteins; Related Reagents and Methods
;  FILE REFERENCE: DX0884K(2d)
;  CURRENT APPLICATION NUMBER: US/09/372,348A
;  CURRENT FILING DATE: 1999-08-11
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
;   LENGTH: 160
;   TYPE: PRT
;   ORGANISM: aves
US-09-372-348-29

  Query Match             39.0%;  Score 336;  DB 10;  Length 160;
  Best Local Similarity   44.4%;  Pred. No. 2.8e-31;
  Matches   59;  Conservative   26;  Mismatches   48;  Indels    0;  Gaps    0;

Qy         27 WPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSKQLKSNINIRFGIILREDIKELFLDLA 86
              |||| || :  ||: |:|||| ||| ||:::||       :||  ::||: ||||:  :
Db         22 WPTHTVCKEENLEIYYKSCDPQQDFAFSIDRCSDVTTHTFDIRAAMVLRQSIKELYAKVD 81

Qy         87 LMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQIYYAGPVNNPEFTIPQGEYQVLLELYT 146
              |:  | :||::|  :|   | |  |||::||| :|| ||:       ||| :| :   |
Db         82 LIINGKTVLSYSETLCGPGLSKLIFCGKKKGEHLYYEGPITLGIKEIPQRDYTITARLTN 141
```

```
Qy        147 EKRSTVACANATI 159
              | |:|||||: |:
Db        142 EDRATVACADFTV 154


RESULT 8
US-10-000-066-1
; Sequence 1, Application US/10000066
; Publication No. US20020086364A1
; GENERAL INFORMATION:
;  APPLICANT: HONJO, TASUKU
;  APPLICANT:  KATO, KEIZO
;  APPLICANT:  TADA, HIDEAKI
;  TITLE OF INVENTION: POLYPEPTIDE, cDNA ENCODING THE SAME, AND USE OF THEM
;  FILE REFERENCE: Q58771
;  CURRENT APPLICATION NUMBER: US/10/000,066
;  CURRENT FILING DATE:  2001-12-04
;  PRIOR APPLICATION NUMBER: 09/529,064
;  PRIOR FILING DATE: 2000-04-17
;  PRIOR APPLICATION NUMBER: PCT/JP98/04515
;  PRIOR FILING DATE: 1998-10-06
;  PRIOR APPLICATION NUMBER: HEI-9-274673
;  PRIOR FILING DATE: 1997-10-07
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 160
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-000-066-1

  Query Match             12.3%;  Score 106;  DB 12;  Length 160;
  Best Local Similarity   26.3%;  Pred. No. 0.00044;
  Matches   36;  Conservative   24;  Mismatches   61;  Indels   16;  Gaps    5;

Qy         33 CSDSGLEVLYQSCDPLQ-DFGFSVEKCSKQLKSNINIRFGIILREDIKELFLDLALMSQG 91
              |: |   : |  || |:     | | | :   :|  :    | | ::| |: :| :
Db         25 CNSSDAIISYSYCDHLKFPISISSEPCIRLRGTNGFVHVEFIPRGNLKYLYFNLFI---- 80

Qy         92 SSVLNFSYP-----ICEAALPKFSFCGRRKGEQIYYAGPVNNPEFTIPQGEYQVLLELY- 145
               || :   |     :|      :|||   ||| :  : | :      |:| |: : |
Db         81 -SVNSIELPKRKEVLCHGHDDDYSFCRALKGETVNTSIPFSFEGILFPKGHYRCVAEAIA 139

Qy        146 --TEKRSTVACANATIM 160
                ||::  : | | ||:
Db        140 GDTEEK--LFCLNFTII 154


RESULT 9
US-09-372-348-23
; Sequence 23, Application US/09372348A
; Publication No. US20030028005A1
; GENERAL INFORMATION:
;  APPLICANT: Hedrick, Joseph A.
;  APPLICANT: Wang, Luquan
;  APPLICANT: Zlotnik, Albert
;  APPLICANT: Murgolo, Nicholas J.
;  APPLICANT: Greene, Jonathan R.
;  APPLICANT: Johnston, James A.
;  APPLICANT: Bazan, J. Fernando
;  TITLE OF INVENTION: Mammalian Proteins; Related Reagents and Methods
;  FILE REFERENCE: DX0884K(2d)
;  CURRENT APPLICATION NUMBER: US/09/372,348A
;  CURRENT FILING DATE: 1999-08-11
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
;   LENGTH: 160
;   TYPE: PRT
;   ORGANISM: primate
US-09-372-348-23

  Query Match             11.8%;  Score 101.5;  DB 10;  Length 160;
  Best Local Similarity   24.1%;  Pred. No. 0.0015;
  Matches   34;  Conservative   26;  Mismatches   70;  Indels   11;  Gaps    3;

Qy          4 FTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQ-DFGFSVEKCSKQL 62
              | :|||  : ||          :|   : ||: |   : |  || :|     :|  | :
Db          6 FFSTLF--SSIFT--------EAQKQYWVCNSSDASISYTYCDKMQYPISINVNPCIELK 55

Qy         63 KSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQIYY 122
               |   :    | | |:|:|: :| :     ::       ||  :   :|||   ||| :
Db         56 GSKGLLHIFYIPRRDLKQLYFNLYITVNTMNLPKRKEVICRGSDDDYSFCRALKGETVNT 115

Qy        123 AGPVNNPEFTIPQGEYQVLLE 143
                    :|:|: ::|
Db        116 TISFSFKGIKFSKGKYKCVVE 136


RESULT 10
US-10-000-066-7
; Sequence 7, Application US/10000066
; Publication No. US20020086364A1
; GENERAL INFORMATION:
;  APPLICANT: HONJO, TASUKU
;  APPLICANT:  KATO, KEIZO
;  APPLICANT:  TADA, HIDEAKI
;  TITLE OF INVENTION: POLYPEPTIDE, cDNA ENCODING THE SAME, AND USE OF THEM
;  FILE REFERENCE: Q58771
;  CURRENT APPLICATION NUMBER: US/10/000,066
;  CURRENT FILING DATE:  2001-12-04
;  PRIOR APPLICATION NUMBER: 09/529,064
;  PRIOR FILING DATE: 2000-04-17
;  PRIOR APPLICATION NUMBER: PCT/JP98/04515
;  PRIOR FILING DATE: 1998-10-06
;  PRIOR APPLICATION NUMBER: HEI-9-274673
;  PRIOR FILING DATE: 1997-10-07
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 160
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-000-066-7

  Query Match             11.8%;  Score 101.5;  DB 12;  Length 160;
  Best Local Similarity   24.1%;  Pred. No. 0.0015;
  Matches   34;  Conservative   26;  Mismatches   70;  Indels   11;  Gaps    3;

Qy          4 FTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQ-DFGFSVEKCSKQL 62
              | :|||  : ||          :|   : ||: |   : |  || :|     :|  | :
Db          6 FFSTLF--SSIFT--------EAQKQYWVCNSSDASISYTYCDKMQYPISINVNPCIELK 55

Qy         63 KSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQIYY 122
               |   :    | | |:|:|: :| :     ::       ||  :   :|||   ||| :
Db         56 GSKGLLHIFYIPRRDLKQLYFNLYITVNTMNLPKRKEVICRGSDDDYSFCRALKGETVNT 115

Qy        123 AGPVNNPEFTIPQGEYQVLLE 143
                    :|:|: ::|
Db        116 TISFSFKGIKFSKGKYKCVVE 136


RESULT 11
US-09-372-348-26
; Sequence 26, Application US/09372348A
; Publication No. US20030028005A1
; GENERAL INFORMATION:
;  APPLICANT: Hedrick, Joseph A.
;  APPLICANT: Wang, Luquan
;  APPLICANT: Zlotnik, Albert
;  APPLICANT: Murgolo, Nicholas J.
;  APPLICANT: Greene, Jonathan R.
;  APPLICANT: Johnston, James A.
;  APPLICANT: Bazan, J. Fernando
;  TITLE OF INVENTION: Mammalian Proteins; Related Reagents and Methods
```

```
;  FILE REFERENCE: DX0884K(2d)
;  CURRENT APPLICATION NUMBER: US/09/372,348A
;  CURRENT FILING DATE: 1999-08-11
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
;   LENGTH: 141
;   TYPE: PRT
;   ORGANISM: primate
US-09-372-348-26

  Query Match             11.2%;  Score 96.5;  DB 10;  Length 141;
  Best Local Similarity   25.6%;  Pred. No. 0.0049;
  Matches   30;  Conservative   20;  Mismatches   56;  Indels   11;  Gaps    3;

Qy         33 CSDSGLEVLYQSCDPLQ-DFGFSVEKCSKQLKSNINIRFGIILREDIKELFLDLALMSQG 91
              |: |   : |  || |:     | | | :   :|  :    | | ::| |: :| :
Db         25 CNSSDAIISYSYCDHLKFPISISSEPCIRLRGTNGFVHVEFIPRGNLKYLYFNLFI---- 80

Qy         92 SSVLNFSYP-----ICEAALPKFSFCGRRKGEQIYYAGPVNNPEFTIPQGEYQVLLE 143
              || :   |     :|      :|||   ||| :  : | :      |:| |: : |
Db         81 -SVNSIELPKRKEVLCHGHDDDYSFCRALKGETVNTSIPFSFEGILFPKGHYRCVAE 136


RESULT 12
US-09-899-917-17
; Sequence 17, Application US/09899917
; Patent No. US20020119552A1
;   GENERAL INFORMATION:
;        APPLICANT: Olsen, Henrik S.
;                   Ruben, Steven M.
;        TITLE OF INVENTION: Human Oncogene Induced Secreted Protein I
;        NUMBER OF SEQUENCES: 17
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Sterne, Kessler Goldstein & Fox P.L.L.C.
;             STREET: 1100 New York Ave., NW, Suite 600
;             CITY: Washington
;             STATE: DC
;             COUNTRY: USA
;             ZIP: 20005-3934
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/899,917
;             FILING DATE: 09-Jul-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/994,962
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: US 60/037,388
;             FILING DATE: 07-FEB-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Steffe, Eric K.
;             REGISTRATION NUMBER: 36,688
;             REFERENCE/DOCKET NUMBER: 1488.0440002
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 202-371-2600
;             TELEFAX: 202-371-2540
;   INFORMATION FOR SEQ ID NO: 17:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 60 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: No. US20020119552A1 Relevant
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-899-917-17

  Query Match             10.9%;  Score 94;  DB 9;  Length 60;
  Best Local Similarity   33.3%;  Pred. No. 0.003;
  Matches   43;  Conservative    3;  Mismatches   13;  Indels   70;  Gaps   11;

Qy         27 WPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSKQLKSNINIRFGIILREDIKELFLDLA 86
              |||||         |  |||| ||| ||   || :|:              ||||
Db          1 WPTHV--------CLEYSCDP-QDF-FS---CSIRLR--------------IKELL---- 29

Qy         87 LMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQIYYAGPVNNPEFTIPQGEYQVLLELYT 146
                                 |    |||  |||   | ||       ||||        |
Db         30 -----------------GVLSCLKFCG--KGE---YYGP-------IPQG--------YL 52

Qy        147 EKRSTVACA 155
              |:  |||||
Db         53 ER--TVACA 59


RESULT 13
US-10-424-599-214951
; Sequence 214951, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 214951
;   LENGTH: 555
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(555)
;   OTHER INFORMATION: unsure at all Xaa locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_36126C.1.pep
US-10-424-599-214951

  Query Match              9.5%;  Score 81.5;  DB 12;  Length 555;
  Best Local Similarity   25.7%;  Pred. No. 1.9;
  Matches   26;  Conservative   18;  Mismatches   32;  Indels   25;  Gaps    4;

Qy         48 LQDFGFSVEKCSKQLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALP 107
              :|  |:  :|| :|::  :  ||  :| | :   | ||                 :|||
Db         66 VQGKGYK-DKCYEQIRKTVEGRFNKLLNELV---FEDL----------------KAALE 104

Qy        108 KFSFCGRRKGEQIYYAGPVNNPEFTIPQGEYQVLLELYTEK 148
              :    |   |:  |  |   | : |    :|::: ||||:
Db        105 EARAIGEELGDVYDYVAPCFPPRYEI----FQLMVNLYTER 141


RESULT 14
US-10-424-599-199134
; Sequence 199134, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 199134
;   LENGTH: 170
;   TYPE: PRT
```

```
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_21842C.1.pep
US-10-424-599-199134

  Query Match             9.4%;  Score 81;  DB 12;  Length 170;
  Best Local Similarity   21.1%;  Pred. No. 0.43;
  Matches   35;  Conservative   20;  Mismatches   59;  Indels   52;  Gaps    7;

Qy         24 GKAW-PTHVVCSDS----GLEVLYQSCDPLQDFGFSVEK---------------------- 57
              | :| |  |    :     |   ||   ||  |:   :||
Db         23 GASWVPNQVTVKEEYKVPGDNPLYFCADP-ADYILQIEKVDLDPNPPKPGEKLAIKASGD 81

Qy         58 ----CSKQLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCG 113
                  :  | :: :::|:|                    :::|     |:       |
Db         82 FKEEVGEGSKMHLQVKYGLI-----------------TLINQERDACDTIEQADLKCP 122

Qy        114 RRKGEQIYYAGPVNNPEFTIPQQGEYQVLLELYTEKRSTVACANATI 159
               :||| :     |  |   || | | || ::|||    : |  | |
Db        123 LKKGE-LSLTKDVELPR-EIPPGTYTVLADVYTEDGDKITCLTAKI 166


RESULT 15
US-10-282-122A-77326
; Sequence 77326, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 77326
;   LENGTH: 876
;   TYPE: PRT
;   ORGANISM: Vibrio cholerae
US-10-282-122A-77326

  Query Match             9.0%;  Score 78;  DB 12;  Length 876;
  Best Local Similarity   26.1%;  Pred. No. 9;
  Matches   37;  Conservative   23;  Mismatches   40;  Indels   42;  Gaps   10;

Qy          8 LFLWTLIFPSCSGGGGGKAWPTHVV--CSDSGLEVLYQSCDPLQDFGFSVEKCSKQLKSN 65
              : :  :: |:||   | :   | |   ||  ||                 |:|    |:
Db        599 IVMTNILCPTCSRPMGIRTASTGVFLGCSGYGLP--------------PKERC----KTT 640

Qy         66 INI--RFGII--LREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQIY 121
              ||:    |:|  | ||::      || ::          |||| |:  :    :||   ::
Db        641 INLGDEEGVINVLEEDVE----TAALRAKKRC------PICETAMDAYLIDDKRK---LH 687

Qy        122 YAGPVNNPE---FTIPQGEYQV 140
                 |  |||    | : :||::|
Db        688 VCG--NNPNCEGPIVEEGEPKV 707


Search completed: April 29, 2004, 15:29:31
Job time : 43 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: April 29, 2004, 15:21:06 ; Search time 21 Seconds
(without alignments)
742.049 Million cell updates/sec

Title: US-09-899-917-2
Perfect score: 862
Sequence: 1 MKGFTATLFLWTLIFPSCSG.....ELYTEKRSTVACANATIMCS 162

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	344	39.9	160	2	S18633	MD-1 protein - chi
2	99	11.5	1137	2	G84581	copia-like retroel
3	81	9.4	406	2	H71139	hypothetical prote
4	79.5	9.2	484	2	H84723	probable triacylgl
5	78	9.0	876	2	B82163	DNA topoisomerase
6	78	9.0	1133	2	S54496	probable membrane
7	77.5	9.0	394	2	A75012	conserved hypothet
8	77.5	9.0	428	2	T14529	S-locus-specific g
9	77	8.9	246	2	AH0393	probable exported
10	77	8.9	739	2	G96740	hypothetical prote
11	76.5	8.9	294	2	T22058	hypothetical prote
12	76.5	8.9	436	2	T28066	hypothetical prote
13	76.5	8.9	452	1	VGBEG3	glycoprotein D pre
14	76.5	8.9	471	2	E86192	hypothetical prote
15	75.5	8.8	342	2	D81306	probable UDP-N-ace
16	75.5	8.8	1215	2	C84848	hypothetical prote
17	75.5	8.8	1241	2	F97286	DNA-dependent RNA
18	75	8.7	699	2	A96529	hypothetical prote
19	74.5	8.6	484	2	T16695	hypothetical prote
20	73.5	8.5	442	1	VGBEEA	glycoprotein D pre
21	73.5	8.5	481	2	B97348	hypothetical prote
22	73	8.5	1106	2	A97819	isoleucine-tRNA li
23	72	8.4	409	2	E90299	conserved hypothet
24	72	8.4	686	1	D70172	DNA recombinase (r
25	72	8.4	1218	2	S71376	glutamate receptor
26	71.5	8.3	267	2	B90486	ABC transporter, A
27	71.5	8.3	1513	1	RNRZC2	DNA-directed RNA p
28	71	8.2	668	2	S56909	polymyxin B resist
29	71	8.2	887	2	T01113	translation initia
30	71	8.2	1213	2	D86384	unknown protein [i
31	70.5	8.2	218	2	G71875	hypothetical prote
32	70.5	8.2	456	2	H71733	outer membrane pro
33	70.5	8.2	532	2	T49873	calnexin homolog -
34	70.5	8.2	566	1	HMIVD1	hemagglutinin prec
35	70.5	8.2	1474	2	D88550	protein ZC84.6 [im
36	70	8.1	175	1	G69518	isochorismatase (e
37	70	8.1	215	2	F82062	adenylylsulfate ki
38	70	8.1	515	2	D84142	1-pyrroline-5-carb
39	70	8.1	794	2	T46073	hypothetical prote
40	70	8.1	950	1	PWBYR1	Ca2+-transporting
41	70	8.1	1116	2	T42213	m-tomosyn, isoform
42	69.5	8.1	440	2	AC0962	D-serine dehydrata
43	69	8.0	316	2	F70441	capsular polysacch
44	69	8.0	385	2	T01332	hypothetical prote
45	69	8.0	591	2	C84220	propionyl-CoA carb

ALIGNMENTS

RESULT 1
S18633
MD-1 protein - chicken
C;Species: Gallus gallus (chicken)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Sep-1999
C;Accession: S18633; S42854
R;Burk, O.; Klempnauer, K.H.
EMBO J. 10, 3713-3719, 1991
A;Title: Estrogen-dependent alterations in differentiation state of myeloid cells caused
A;Reference number: S18633; MUID:92037533; PMID:1718743
A;Accession: S18633
A;Molecule type: mRNA
A;Residues: 1-160 <BUR>
A;Cross-references: EMBL:X60450
R;Burk, O.
submitted to the EMBL Data Library, September 1991
A;Reference number: S42854
A;Accession: S42854
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-130,'R',132-160 <BU2>
A;Cross-references: EMBL:X60450; NID:g432637; PID:g432638
C;Superfamily: chicken MD-1 protein

Query Match 39.9%; Score 344; DB 2; Length 160;
Best Local Similarity 45.1%; Pred. No. 2.6e-27;
Matches 60; Conservative 26; Mismatches 47; Indels 0; Gaps 0;

```
Qy         27 WPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSKQLKSNINIRFGIILREDIKELFLDLA 86
              |||| || :  ||: |:|||| ||| ||:::||      :||  ::||: ||||:  :
Db         22 WPTHTVCKEENLEIYYKSCDPQQDFAFSIDRCSDVTTHTFDIRAAMVLRQSIKELYAKVD 81

Qy         87 LMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQIYYAGPVNNPEFTIPQGEYQVLLELYT 146
              |:  | :||::|  :|   | |  |||::||| :|| ||:      ||||:| :   |
Db         82 LIINGKTVLSYSETLCGPGLSKLIFCGKKKGEHLYYEGPITLGIKEIPQGDYTITARLTN 141

Qy        147 EKRSTVACANATI 159
              | |:|||||: |:
Db        142 EDRATVACADFTV 154
```

RESULT 2
G84581
copia-like retroelement pol polyprotein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: G84581
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: G84581
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1137 <STO>
A;Cross-references: GB:AE002093; NID:g3687234; PIDN:AAC62132.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g19840
A;Map position: 2

Query Match 11.5%; Score 99; DB 2; Length 1137;
Best Local Similarity 26.3%; Pred. No. 0.13;
Matches 36; Conservative 23; Mismatches 42; Indels 36; Gaps 7;

```
Qy         10 LWTLIFPSCSG-----------GGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSV--- 55
Db        448 LWTSAVPNFSGLKRFGCVVYVYSQEGKLDPR----AKKGVFVGY----PNGVKGFRVWMI 499

Qy         56 --EKCSKQLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCG 113
Db        500 EEERCS--------ISRNVVFREDV--MYKD--ILNQSTSGMSFDFPLATNRIPSFECAG 547

Qy        114 RRKGEQIYYAGPVNNPE 130
Db        548 NRKEDEISVQGGVSDDD 564
```

RESULT 3
H71139
hypothetical protein PH0331 - Pyrococcus horikoshii
C;Species: Pyrococcus horikoshii
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 28-Jul-2000
C;Accession: H71139
R;Kawarabayasi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A;Reference number: A71000; MUID:98344137; PMID:9679194
A;Accession: H71139
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-406 <KAW>
A;Cross-references: GB:AP000002; NID:g3236129; PIDN:BAA29405.1; PID:g3256722
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank
C;Genetics:
A;Gene: PH0331
C;Superfamily: conserved hypothetical protein PAB1277

Query Match 9.4%; Score 81; DB 2; Length 406;
Best Local Similarity 25.8%; Pred. No. 2.6;
Matches 39; Conservative 27; Mismatches 49; Indels 36; Gaps 9;

```
Qy          1 MKG--FTATLFLWTLIFPSCSGGGGGKA-WPTHVVCSDSGLEVLYQSCDPLQDFGF--SV 55
Db          4 MKGASFFVTLFLWFIVMSLIFGVPGKLAVFPLLLLI----MGLLF---DAPGDFEVEREI 56

Qy         56 EKC----SKQLKSNINIRFG-----IILREDIKELFLDLALMSQGSSVLNFSYPICEAAL 106
Db         57 EKAQTFVGNEIEVLVRVRVGRGIGLVMVRENIPKAF----MTSSGSNVGYF--------- 103

Qy        107 PKFSFCGRRKGEQIYYAGPVNNPEFTIPQGE 137
Db        104 --FTYPGRRSFQLSYKLLPLKRGVYEIPKTE 132
```

RESULT 4
H84723
probable triacylglycerol lipase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: H84723
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: H84723
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-484 <STO>
A;Cross-references: GB:AE002093; NID:g4582461; PIDN:AAD24845.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g31690
A;Map position: 2

Query Match 9.2%; Score 79.5; DB 2; Length 484;
Best Local Similarity 27.4%; Pred. No. 4.6;
Matches 34; Conservative 18; Mismatches 43; Indels 29; Gaps 5;

```
Qy         35 DSGLEVLYQSCDPLQDFGFSVEKCSKQLKSNINIRFGIILREDIKELFLDLALMSQGSSV 94
Db        248 ESGFLSLYTSDESESKFG--LESCRQQLLSEIS---RLMNKYKGEEMSITLAGHSMGSSL 302

Qy         95 LN-FSYPICEAALPK-----------FSFCGRRKGE-----------QIYYAGPVNNPE 130
Db        303 AQLLAYDIAELGLNRRIGKGDIPVTVFSFAGPRVGNLEFKKRCEELGVKVLRITNVNDPV 362

Qy        131 FTIP 134
Db        363 TKLP 366
```

RESULT 5
B82163
DNA topoisomerase I VC1730 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C;Accession: B82163
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: B82163
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-876 <HEI>
A;Cross-references: GB:AE004251; GB:AE003852; NID:g9656248; PIDN:AAF94880.1; GSPDB:GN001
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC1730
A;Map position: 1
C;Superfamily: bacterial type I DNA topoisomerase

Query Match 9.0%; Score 78; DB 2; Length 876;
Best Local Similarity 26.1%; Pred. No. 13;
Matches 37; Conservative 23; Mismatches 40; Indels 42; Gaps 10;

```
Qy          8 LFLWTLIFPSCSGGGGGKAWPTHVV--CSDSGLEVLYQSCDPLQDFGFSVEKCSKQLKSN 65
Db        599 IVMTNILCPTCSRPMGIRTASTGVFLGCSGYGLP-------------PKERC----KTT 640

Qy         66 INI--RFGII--LREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQIY 121
Db        641 INLGDEEGVINVLEEDVE----TAALRAKKRC------PICETAMDAYLIDDKRK---LH 687

Qy        122 YAGPVNNPE---FTIPQGEYQV 140
Db        688 VCG--NNPNCEGFIVEEGEPKV 707
```

RESULT 6
S54496

```
probable membrane protein YPR022c - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein YP9367.02c
C;Species: Saccharomyces cerevisiae
C;Date: 08-Jul-1995 #sequence_revision 19-Oct-1995 #text_change 19-Apr-2002
C;Accession: S54496
R;Badcock, K.; Churcher, C.M.
submitted to the EMBL Data Library, May 1995
A;Reference number: S54059
A;Accession: S54496
A;Molecule type: DNA
A;Residues: 1-1133 <BAD>
A;Cross-references: EMBL:Z49274; NID:g809585; PID:g809587; GSPDB:GN00016; MIPS:YPR022c
A;Experimental source: strain AB972
C;Genetics:
A;Gene: MIPS:YPR022c
A;Cross-references: SGD:S0006226
A;Map position: 16R
C;Keywords: transmembrane protein
F;532-548/Domain: transmembrane #status predicted <TM1>
F;768-784/Domain: transmembrane #status predicted <TM2>
F;931-947/Domain: transmembrane #status predicted <TM3>
F;971-987/Domain: transmembrane #status predicted <TM4>

  Query Match             9.0%;  Score 78;  DB 2;  Length 1133;
  Best Local Similarity   25.3%;  Pred. No. 17;
  Matches   37;  Conservative   16;  Mismatches   31;  Indels   62;  Gaps    8;

Qy          46 DPLQDFGFSVEKCS-----------------KQLKS---NINIRFGII----LREDIKE 80
               |||  :  | |||                  | |||   | : ::|::       | :
Db         472 DPLTKYKISKEKCQEMFSFVPELRYVSIESIHKSLKSFWLNFHPQYGLLHKPSFHVDKQP 531

Qy          81 LFLDLALMSQGSSVLNFSY------PICEAALPKFSFCGRRKGEQIYYAGPV-----NNP 129
                 |:|||:  |:| |   |       |||                    ||:     ::
Db         532 AILNLALIMTGASFLGSEYREQISDPIC--------------------GPLRWIIFSHA 570

Qy         130 EFTIPQGEYQV----LLELYTEKRST 151
               :|  |   | :    |:| | || ||
Db         571 DFQPPSKTYIIQSLLLVEGY-EKTST 595


RESULT 7
A75012
conserved hypothetical protein PAB1277 - Pyrococcus abyssi (strain Orsay)
C;Species: Pyrococcus abyssi
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 28-Jul-2000
C;Accession: A75012
R;anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A;Reference number: A75001
A;Accession: A75012
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-394 <KAW>
A;Cross-references: GB:AJ248288; GB:AL096836; NID:g5458960; PIDN:CAB50535.1; PID:g545904
A;Experimental source: strain Orsay
C;Genetics:
A;Gene: PAB1277
C;Superfamily: conserved hypothetical protein PAB1277

  Query Match             9.0%;  Score 77.5;  DB 2;  Length 394;
  Best Local Similarity   23.4%;  Pred. No. 5.7;
  Matches   36;  Conservative   31;  Mismatches   44;  Indels   43;  Gaps    9;

Qy           1 MKG--FTATLFLWTLIFPSCSGGGGGKA-WPTHVVCSDSGLEVLYQSCDPLQDFGFSVEK 57
               |:|  |  ||||| ::     |  | | :|  ::     : ::: :       ||:||:
Db           1 MRGVSFFLTLFLWFVLLSVIFGTPGKLALFPLLIII----VGMIFDT-----PGGFTVER 51

Qy          58 --------CSKQLKSNINIRFG-----IILREDIKELFLDLALMSQGSSVLN-FSYPICE 103
                       :::: : :|       :||:|::   |       :||:  : |:||
Db          52 EVTKRRVFVGEEVRVKLRVRVSKGIGIVILKENLPRAF-----EVKGSATFHFFTYPGKR 106

Qy         104 AALPKFSFCGRRKGEQIYYAGPVNNPEFTIPQGE 137
                 ::||  ||||           |: ||: |
Db         107 EFSFEYSFVPRRKG-----------EYDIPKTE 128


RESULT 8
T14529
S-locus-specific glycoprotein - wild cabbage (fragment)
N;Alternate names: S glycoprotein
C;Species: Brassica oleracea (wild cabbage)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 03-Nov-2000
C;Accession: T14529
R;Kusaba, M.; Nishio, T.; Satta, Y.; Hinata, K.; Ockendon, D.
Proc. Natl. Acad. Sci. U.S.A. 94, 7673-7678, 1997
A;Title: Striking sequence similarity in inter- and intra-specific comparisons o f class
echanism.
A;Reference number: Z18078; MUID:97352858; PMID:9207151
A;Accession: T14529
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-428 <KUS>
A;Cross-references: EMBL:D85205; NID:g2351141; PIDN:BAA21939.1; PID:g2351142
C;Superfamily: S-locus-specific glycoprotein; S-locus-specific glycoprotein homology
C;Keywords: glycoprotein

  Query Match             9.0%;  Score 77.5;  DB 2;  Length 428;
  Best Local Similarity   25.2%;  Pred. No. 6.3;
  Matches   27;  Conservative   20;  Mismatches   49;  Indels   11;  Gaps    4;

Qy          49 QDFGFSVEKCSKQLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPK 108
               | | |  :    ::|   |:: |:           ||     |   |  ::|| :    ||:
Db         149 QSFDFPTDTLLPEMKLGYNLKTGL-------NRFLTAWRNSDDPSSGDYSYKLENRELPE 201

Qy         109 FSFCGRRKGEQIYYAGPVNNPEFT-IPQGE-YQVLLELYTEKRSTVA 153
               |     : | |:: :|| |   |: ||: :    ::  :||    ||
Db         202 FYLL--KSGFQVHRSGPWNGVRFSGIPENQKLSYMVYNFTENSEEVA 246


RESULT 9
AH0393
probable exported protein YPO3241 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
C;Accession: AH0393
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; E
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, E
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AH0393
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-246 <KUR>
A;Cross-references: GB:AL590842; PIDN:CAC92476.1; PID:g15981177; GSPDB:GN00175
C;Genetics:
A;Gene: YPO3241
C;Superfamily: Escherichia coli probable membrane protein yafK

  Query Match             8.9%;  Score 77;  DB 2;  Length 246;
  Best Local Similarity   29.0%;  Pred. No. 3.8;
  Matches   31;  Conservative   12;  Mismatches   30;  Indels   34;  Gaps    6;

Qy          45 CDPLQDFGFSVEK---CSKQLKSNI---NIRFGIILREDIKELFLDLALMSQGSSVLNFS 98
               | |   |  |  :    ||:||  :   ::    |   | : ||:  |    ||   | |
Db          13 CLPAASFTSSASEPVAVSKELKQQLLGSSVYIQIFKEERVLELYAKL----QGEYRLVQS 68

Qy          99 YPICEAALPKFSFCG-----RRKGEQIYYAGPVNNPEFTIPQGEYQV 140
               |||||       | |     ||:|             :|  |:| |::
Db          69 YPICE-------FSGGLGPKRREG-----------DFKSPEGFYRI 96


RESULT 10
G96740
```

hypothetical protein F14O23.20 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: G96740
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: G96740
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-739 <STO>
A;Cross-references: GB:AE005173; NID:g7239509; PIDN:AAF43235.1; GSPDB:GN00141
C;Genetics:
A;Gene: F14O23.20
A;Map position: 1

Query Match 8.9%; Score 77; DB 2; Length 739;
Best Local Similarity 24.0%; Pred. No. 13;
Matches 25; Conservative 17; Mismatches 30; Indels 32; Gaps 4;

```
Qy         48 LQDFGFSVEKCSKQLKSNINIRFGIILR---EDIKELFLDLALMSQGSSVLNFSYPICEA 104
              :|  |:  :|| :|::  :  ||  :|    ||:|                       |
Db        264 VQGKGYK-DKCYEQIRKAVEDRFNRLLTLVFEDLK-----------------------A 298

Qy        105 ALPKFSFCGRRKGEQIYYAGPVNNPEFTIPQGEYQVLLELYTEK 148
              || :    |   |:   | |   | : |    :|::: ||||:
Db        299 ALEEARMIGEELGDIYDYVAPCFPPRYEI----FQLMVNLYTER 338
```

RESULT 11
T22058
hypothetical protein F40G12.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Sep-2000
C;Accession: T22058
R;McMurray, A.
submitted to the EMBL Data Library, July 1996
A;Reference number: Z19507
A;Accession: T22058
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-294 <WIL>
A;Cross-references: EMBL:Z77661; PIDN:CAB01187.1; GSPDB:GN00023; CESP:F40G12.4
A;Experimental source: clone F40G12
C;Genetics:
A;Gene: CESP:F40G12.4
A;Map position: 5
A;Introns: 59/3; 159/3; 262/1
C;Superfamily: Caenorhabditis elegans hypothetical protein F40G12.4

Query Match 8.9%; Score 76.5; DB 2; Length 294;
Best Local Similarity 26.7%; Pred. No. 5.2;
Matches 23; Conservative 13; Mismatches 49; Indels 1; Gaps 1;

```
Qy         28 PTHVVCSDSGLEVLYQSCDPLQD-FGFSVEKCSKQLKSNINIRFGIILREDIKELFLDLA 86
              | |  |||:   :  | ||  :    : :||  ::|    |    | :  :|||   ||
Db        107 PAHGGCSDTYYRLNAQRCDAQKTMMEWEIEKQLERLPKFKNSTQVIAMCREIKECMEDLC 166

Qy         87 LMSQGSSVLNFSYPICEAALPKFSFC 112
                 : ||  : |  :   |: |
Db        167 FTEDSQYEIEFSLAVPELTVSHFTVC 192
```

RESULT 12
T28066
hypothetical protein ZK867.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C;Accession: T28066
R;Nhan, M.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid ZK867.
A;Reference number: Z20464
A;Accession: T28066
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-436 <NHA>
A;Cross-references: EMBL:U41039; PIDN:AAA82394.1; CESP:ZK867.2
C;Genetics:
A;Gene: CESP:ZK867.2
A;Introns: 47/3; 79/1; 108/3; 162/1; 212/2; 295/3; 340/3; 367/3

Query Match 8.9%; Score 76.5; DB 2; Length 436;
Best Local Similarity 21.0%; Pred. No. 8.1;
Matches 34; Conservative 26; Mismatches 67; Indels 35; Gaps 6;

```
Qy         33 CSDSGLE-----------VLYQSCDPLQDFGFSVEKCSKQLKSNINIRFGIILREDIKE 80
              |  | |             | |   |  : ||   :  |||   | :  |::  : :
Db         24 CKHPGAEVEILKMIFRLIGVNYTMIDVWKKFGQQYDFGSKQKNGNWSGMIGLLQSDQLDM 83

Qy         81 LFLDLALMSQGSSVLNFSYP--ICEAALPKFSFCGRRK----------GEQIYYAG---- 124
              : | : :  :   |: ||||  : | ::  |    |                |:|
Db         84 IGLSMRIAPEREEVVLFSYPTRVFETSIQSFPVSSRMLLLIILIATFFISQLYQTDMLAF 143

Qy        125 ---PVNNPEFTIPQGEYQVLLELYTEKRSTVAC-ANATIMCS 162
                 |:    ::||    :  |||   ::  :|   | |::|:
Db        144 LSVPLT---YSIPFRSIKQALELVEHQKMYIAAFENQTLLCT 182
```

RESULT 13
VGBEG3
glycoprotein D precursor - equine herpesvirus 1
C;Species: equine herpesvirus 1
A;Note: host Equus caballus (domestic horse)
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 16-Jul-1999
C;Accession: I36802; B36646; PQ0146
R;Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
submitted to GenBank, March 1992
A;Description: The DNA sequence of equine herpesvirus-1.
A;Reference number: A36805
A;Accession: I36802
A;Molecule type: DNA
A;Residues: 1-452 <TEL>
A;Cross-references: GB:M86664; NID:g330791; PIDN:AAB02507.1; PID:g330863
A;Experimental source: strain Ab4p
R;Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
Virology 189, 304-316, 1992
A;Title: The DNA sequence of equine herpesvirus-1.
A;Reference number: A41831; MUID:92295566; PMID:1318606
A;Contents: annotation; possible protein-coding frames
A;Note: neither amino acid nor nucleotide sequence is given
R;Audonnet, J.C.; Winslow, J.; Allen, G.; Paoletti, E.
J. Gen. Virol. 71, 2969-2978, 1990
A;Title: Equine herpesvirus type 1 unique short fragment encodes glycoproteins with homol
A;Reference number: A36646; MUID:91108393; PMID:2177089
A;Accession: B36646
A;Molecule type: DNA
A;Residues: 1-452 <AUD>
A;Cross-references: GB:M86664; NID:g330791; PIDN:AAB02507.1; PID:g330863
A;Experimental source: strain Kentucky D
R;Elton, D.M.; Halliburton, I.W.; Killington, R.A.; Meredith, D.M.; Bonass, W.A.
Gene 101, 203-208, 1991
A;Title: Sequence analysis of the 4.7-kb BamHI-EcoRI fragment of the equine herpesvirus t
A;Reference number: JQ0998; MUID:91276272; PMID:1647359
A;Accession: PQ0146
A;Molecule type: DNA
A;Residues: 292-452 <ELT>
A;Cross-references: GB:M36299; NID:g330787; PIDN:AAA66546.1; PID:g808672

```
C;Genetics:
A;Gene: 72
C;Superfamily: herpesvirus glycoprotein D
C;Keywords: glycoprotein; transmembrane protein
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-452/Product: glycoprotein D #status predicted <MAT>
F;399-419/Domain: hydrophobic <HYD>
F;404-422/Domain: transmembrane #status predicted <TM2>
F;103,111,347,396/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;347,396/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             8.9%;  Score 76.5;  DB 1;  Length 452;
  Best Local Similarity   26.0%;  Pred. No. 8.5;
  Matches   44;  Conservative   16;  Mismatches   62;  Indels   47;  Gaps    8;

Qy          2 KGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSKQ 61
              | ::| |  |  | |:|        | | |        :| |  |:|   |    |
Db        161 KTYSARL-TWFKIMPTC-------ATPIH--------DVSYMKCNPKLSFAMCDE----- 199

Qy         62 LKSNINIRFGII--LREDIKELFLDLALMSQGSSVL-------------NFSYPICEAA 105
              :|:|  :  :|    |  || | ||  :  :| |                :||  |
Db        200 -RSDILWQASLITMAAETDDELGLVLAAPAHSASGLYRRVIEIDGRRIYTDFSVTIPSER 258

Qy        106 LP-----KFSFCGRRKGEQIYYAGPVNN----PEFTIPQGEYQVLLELY 145
               |       |    | |  : |  | |       ||  ||||:   |: :
Db        259 CPIAFEQNFGNPDRCKTPEQYSRGEVFTRRFLGEFNFPQGEHMTWLKFW 307


RESULT 14
E86192
hypothetical protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: E86192
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: E86192
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-471 <STO>
A;Cross-references: GB:AE005172; NID:g6850308; PIDN:AAF29385.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1

  Query Match             8.9%;  Score 76.5;  DB 2;  Length 471;
  Best Local Similarity   26.6%;  Pred. No. 8.9;
  Matches   33;  Conservative   19;  Mismatches   43;  Indels   29;  Gaps    5;

Qy         35 DSGLEVLYQSCDPLQDFGFSVEKCSKQLKSNINIRFGIILREDIKELFLDLALMSQGSSV 94
              :||   || | :    ||  :| | :|| | |:    :: :   :|: : ||  | |||:
Db        237 ESGFLGLYTSGESESKFG--LESCREQLLSEIS---RLMNKHKGEEISITLAGHSMGSSL 291

Qy         95 LN-FSYPICEAALPK-----------FSFCGRRKGE-----------QIYYAGPVNNPE 130
                 :| | |  : :            |||| | | |             ::    ||:|
Db        292 AQLLAYDIAELGMNQRRDEKPVPVTVFSFAGPRVGNLGFKKRCEELGVKVLRITNVNDPI 351

Qy        131 FTIP 134
                :|
Db        352 TKLP 355


RESULT 15
D81306
probable UDP-N-acetylglucosamine-N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecapre
C;Species: Campylobacter jejuni
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C;Accession: D81306
R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell
Nature 403, 665-668, 2000
A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hype
A;Reference number: A81250; MUID:20150912; PMID:10688204
A;Accession: D81306
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-342 <PAR>
A;Cross-references: GB:AL139077; GB:AL111168; NID:g6968444; PIDN:CAB73295.1; PID:g6968475
A;Experimental source: serotype O2, strain NCTC 11168
C;Genetics:
A;Gene: murG; Cj1039
C;Superfamily: murG protein
C;Keywords: glycosyltransferase; hexosyltransferase

  Query Match             8.8%;  Score 75.5;  DB 2;  Length 342;
  Best Local Similarity   25.7%;  Pred. No. 7.8;
  Matches   37;  Conservative   22;  Mismatches   34;  Indels   51;  Gaps   11;

Qy         35 DSGLEVLYQSCDPLQDFGFSVEKCSKQLKSNINIR-----FGIILREDIKELFLDLALMS 89
              :  :::::| |     ||    ||| |  :| :||:     | : | | :|    |||:
Db        196 EQNIKIIHQ-CGK-NDF----EKCKKHYQS-LNIQADIFDFSLNLEEKMKN--ADLAISR 246

Qy         90 QGSSVLNFSYPICEAALP-----------------KF----SFCGRRKGEQIYYAGPVN 127
              |:| |   : :|   ||                 ||    : |       ||:    :|
Db        247 AGASTL---FELCANTLPTIFIPYPYAAKNHQYFNAKFLQDQALC------QIFMQNSIN 297

Qy        128 NPEFTIPQGEYQVLLELYTEKRST 151
              ||      :: :|:|  |  ||
Db        298 LDEF------FKSILKLNLENIST 315


Search completed: April 29, 2004, 15:24:22
Job time : 23 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: April 29, 2004, 15:21:06 ; Search time 17 Seconds
(without alignments)
496.198 Million cell updates/sec

Title: US-09-899-917-2
Perfect score: 862
Sequence: 1 MKGFTATLFLWTLIFPSCSG.....ELYTEKRSTVACANATIMCS 162

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	862	100.0	162	1	LY86_HUMAN	O95711 homo sapien
2	613	71.1	162	1	LY86_MOUSE	O88188 mus musculu
3	336	39.0	160	1	LY86_CHICK	Q90890 gallus gall
4	125	14.5	160	1	LY96_BOVIN	P58754 bos taurus
5	124	14.4	160	1	LY96_CRIGR	P58755 cricetulus
6	106	12.3	160	1	LY96_MOUSE	Q9jhf9 mus musculu
7	101.5	11.8	160	1	LY96_HUMAN	Q9y6y9 homo sapien
8	78	9.0	876	1	TOP1_VIBCH	Q9krb2 vibrio chol
9	77	8.9	205	1	CYSC_VIBPA	Q87sx6 vibrio para
10	77	8.9	207	1	CYSC_VIBVU	Q8de75 vibrio vuln
11	76.5	8.9	402	1	VGLD_HSVEA	P24872 equine herp
12	76.5	8.9	452	1	VGLD_HSVEB	P24379 equine herp
13	75.5	8.8	342	1	MURG_CAMJE	Q9pnq2 campylobact
14	75.5	8.8	1241	1	RPOB_CLOAB	Q97eg9 clostridium
15	75	8.7	522	1	UAP1_HUMAN	Q16222 h udp-n-ace
16	73.5	8.5	442	1	VGLD_HSVEK	P22484 equine herp
17	73.5	8.5	488	1	GATA_THETN	Q8rc40 thermoanaer
18	72.5	8.4	892	1	YB96_HUMAN	Q96km6 homo sapien
19	72	8.4	686	1	RECG_BORBU	O51528 borrelia bu
20	71.5	8.3	214	1	RPOA_EUGVI	Q8sl86 euglena vir
21	71.5	8.3	1513	1	RPOD_ORYSA	P12093 oryza sativ
22	71	8.2	668	1	PBS2_YEAST	P08018 saccharomyc
23	70.5	8.2	532	1	CAX2_ARATH	Q38798 arabidopsis
24	70	8.1	215	1	CYSC_VIBCH	Q9kp21 vibrio chol
25	70	8.1	515	1	ROC2_BACHD	Q9k5z5 bacillus ha
26	70	8.1	950	1	ATC1_YEAST	P13586 saccharomyc
27	69.5	8.1	440	1	SDHD_SALTI	Q8z214 salmonella
28	69.5	8.1	440	1	SDHD_SALTY	Q8zl08 salmonella
29	69	8.0	708	1	NICA_MOUSE	P57716 mus musculu
30	69	8.0	7180	1	R1AB_CVMJH	P19751 m replicase
31	68.5	7.9	200	1	RUVA_STAAM	Q99tl1 staphylococ
32	68.5	7.9	474	1	PPAR_XENLA	P37232 xenopus lae
33	68.5	7.9	1616	1	RRPO_TMOB	P90211 tobamovirus
34	68	7.9	387	1	PEP3_RABIT	P27822 oryctolagus
35	68	7.9	458	1	FKB4_HUMAN	Q02790 homo sapien
36	68	7.9	708	1	NICA_RAT	Q8cgu6 rattus norv
37	68	7.9	849	1	SRK6_BRAOL	Q09092 brassica ol
38	67.5	7.8	200	1	RUVA_STAEP	Q8cs90 staphylococ
39	67.5	7.8	422	1	TRB1_SULTO	Q971z5 sulfolobus
40	67.5	7.8	434	1	Y610_METJA	Q58027 methanococc
41	67.5	7.8	457	1	FKB4_RABIT	P27124 oryctolagus
42	67.5	7.8	566	1	HEMA_IADA4	P26562 influenza a
43	67.5	7.8	627	1	K2C1_MOUSE	P04104 mus musculu
44	67	7.8	1422	1	PTPG_CHICK	Q98936 gallus gall
45	67	7.8	1914	1	STCK_EMENI	Q00706 emericella

ALIGNMENTS

bad date

RESULT 1
LY86_HUMAN
ID LY86_HUMAN STANDARD; PRT; 162 AA.
AC O95711; Q9UQC4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Lymphocyte antigen 86 precursor (MD-1 protein).
GN LY86 OR MD1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal liver, and Spleen;
RX MEDLINE=98438337; PubMed=9763566;
RA Miura Y., Shimazu R., Miyake K., Akashi S., Ogata H., Yamashita Y.,
RA Narisawa Y., Kimoto M.;
RT "RP105 is associated with MD-1 and transmits an activation signal in
RT human B cells.";
RL Blood 92:2815-2822(1998).
RN [2]
RP SEQUENCE OF 45-162 FROM N.A.
RC TISSUE=Monocytes;
RX MEDLINE=99180501; PubMed=10079183;
RA Begum N.A., Tsuji S., Nomura M., Shida K., Azuma I., Hayashi A.,
RA Matsumoto M., Seya T., Toyoshima K.;
RT "Human MD-1 homologue is a BCG-regulated gene product in monocytes:
RT Its identification by differential display.";
RL Biochem. Biophys. Res. Commun. 256:325-329(1999).
RN [3]
RP SEQUENCE FROM N.A.
RA Williams S.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

```
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: May cooperate with CD180 and TLR4 to mediate the innate
CC       immune response to bacterial lipopolysaccharide (LPS) and cytokine
CC       production. Important for efficient CD180 cell surface expression
CC       (By similarity).
CC   -!- SUBUNIT: Binds CD180.
CC   -!- SUBCELLULAR LOCATION: Extracellular; associated with CD180 at the
CC       cell surface.
CC   -!- TISSUE SPECIFICITY: Highly expressed in B-cells, monocytes and
CC       tonsil.
CC   -!- INDUCTION: In monocytes, down-regulated by the cell-wall fraction
CC       of Mycobacterium bovis (BCG-CWS).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF057178; AAC98152.2; -.
DR   EMBL; AB020499; BAA76410.1; ALT_INIT.
DR   EMBL; AL031123; CAB76262.1; -.
DR   EMBL; BC038846; AAH38846.1; -.
DR   Genew; HGNC:16837; LY86.
DR   MIM; 605241; -.
DR   GO; GO:0005886; C:plasma membrane; TAS.
DR   GO; GO:0006915; P:apoptosis; TAS.
DR   GO; GO:0008283; P:cell proliferation; TAS.
DR   GO; GO:0006959; P:humoral immune response; TAS.
DR   InterPro; IPR003172; E1_DerP2_DerF2.
DR   InterPro; IPR007110; Ig-like.
DR   Pfam; PF02221; E1_DerP2_DerF2; 1.
DR   SMART; SM00737; ML; 1.
KW   Immune response; Inflammatory response; Signal; Polymorphism.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21    162       LYMPHOCYTE ANTIGEN 86.
FT   CARBOHYD     96     96       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    156    156       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT     160    160       M -> V (in dbSNP:1802323).
FT                                /FTId=VAR_014539.
SQ   SEQUENCE   162 AA;  17905 MW;  3E6497E2DB4C6F27 CRC64;

  Query Match             100.0%;  Score 862;  DB 1;  Length 162;
  Best Local Similarity   100.0%;  Pred. No. 2e-82;
  Matches  162;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MKGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSK 60

Qy         61 QLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 QLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQI 120

Qy        121 YYAGPVNNPEFTIPQGEYQVLLELYTEKRSTVACANATIMCS 162
              |||||||||||||||||||||||||||||||||||||||||
Db        121 YYAGPVNNPEFTIPQGEYQVLLELYTEKRSTVACANATIMCS 162


RESULT 2
LY86_MOUSE
ID   LY86_MOUSE     STANDARD;      PRT;   162 AA.
AC   O88188;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Lymphocyte antigen 86 precursor (MD-1 protein).
GN   LY86 OR MD1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 20-44.
RC   TISSUE=B-cell;
RX   MEDLINE=98349386; PubMed=9686597;
RA   Miyake K., Shimazu R., Kondo J., Niki T., Akashi S., Ogata H.,
RA   Yamashita Y., Miura Y., Kimoto M.;
RT   "Mouse MD-1, a molecule that is physically associated with RP105 and
RT   positively regulates its expression.";
RL   J. Immunol. 161:1348-1353(1998).
RN   [2]
RP   FUNCTION.
RX   MEDLINE=20384784; PubMed=10925274;
RA   Gorczynski R.M., Chen Z., Clark D.A., Hu J., Yu G., Li X., Tsang W.,
RA   Hadidi S.;
RT   "Regulation of gene expression of murine MD-1 regulates subsequent T
RT   cell activation and cytokine production.";
RL   J. Immunol. 165:1925-1932(2000).
CC   -!- FUNCTION: May cooperate with CD180 and TLR4 to mediate the innate
CC       immune response to bacterial lipopolysaccharide (LPS) and cytokine
CC       production. Important for efficient CD180 cell surface expression.
CC   -!- SUBUNIT: Binds CD180.
CC   -!- SUBCELLULAR LOCATION: Extracellular; associated with CD180 at the
CC       cell surface.
CC   -!- TISSUE SPECIFICITY: Highly expressed in spleen, liver, brain and
CC       thymus, and at lower levels in kidney.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB007599; BAA32399.1; -.
DR   MGD; MGI:1321404; Ly86.
DR   InterPro; IPR003172; E1_DerP2_DerF2.
DR   InterPro; IPR007110; Ig-like.
DR   Pfam; PF02221; E1_DerP2_DerF2; 1.
DR   SMART; SM00737; ML; 1.
KW   Immune response; Inflammatory response; Signal.
FT   SIGNAL        1     19
FT   CHAIN        20    162       LYMPHOCYTE ANTIGEN 86.
FT   CARBOHYD     96     96       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    156    156       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   162 AA;  17811 MW;  EED25DEA64A9372E CRC64;

  Query Match             71.1%;  Score 613;  DB 1;  Length 162;
  Best Local Similarity   66.7%;  Pred. No. 1.5e-56;
  Matches  108;  Conservative   22;  Mismatches   32;  Indels    0;  Gaps    0;

Qy          1 MKGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSK 60
              |  |  | |  :|  |  ||  |  |    ||  |  |:  ||||:||||||||||| |:::|||
Db          1 MNGVAAALLVWILTSPSSSDHGSENGWPKHTACNSGGLEVVYQSCDPLQDFGLSIDQCSK 60

Qy         61 QLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQI 120
              |::||:|||||||||:|||::||||: ||::||||:||:|||:||    ||||||||||||||
Db         61 QIQSNLNIRFGIILRQDIRKLFLDITLMAKGSSILNYSYPLCEEDQPKFSFCGRRKGEQI 120

Qy        121 YYAGPVNNPEFTIPQGEYQVLLELYTEKRSTVACANATIMCS 162
              |||||||||   :||||||:||||| | |:|||||||:  |
Db        121 YYAGPVNNPGLDVPQGEYQLLLELYNENRATVACANATVTSS 162


RESULT 3
LY86_CHICK
ID   LY86_CHICK     STANDARD;      PRT;   160 AA.
AC   Q90890;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
```

```
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Lymphocyte antigen 86 precursor (MD-1 protein).
GN   LY86 OR MD1.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Bone marrow;
RX   MEDLINE=92037533; PubMed=1718743;
RA   Burk O., Klempnauer K.-H.;
RT   "Estrogen-dependent alterations in differentiation state of myeloid
RT   cells caused by a v-myb/estrogen receptor fusion protein.";
RL   EMBO J. 10:3713-3719(1991).
CC   -!- FUNCTION: May cooperate with CD180 and TLR4 to mediate the innate
CC       immune response to bacterial lipopolysaccharide (LPS) and cytokine
CC       production. Important for efficient CD180 cell surface expression
CC       (By similarity).
CC   -!- SUBUNIT: Binds CD180 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Extracellular; associated with CD180 at the
CC       cell surface (By similarity).
CC   -!- TISSUE SPECIFICITY: Detected in the macrophage-like 10.4 cells.
CC   -!- INDUCTION: By myb.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X60450; CAA42980.1; -.
DR   PIR; S18633; S18633.
DR   InterPro; IPR003172; E1_DerP2_DerF2.
DR   InterPro; IPR007110; Ig-like.
DR   Pfam; PF02221; E1_DerP2_DerF2; 1.
DR   SMART; SM00737; ML; 1.
KW   Immune response; Inflammatory response; Signal.
FT   SIGNAL       1     20       POTENTIAL.
FT   CHAIN       21    160       LYMPHOCYTE ANTIGEN 86.
SQ   SEQUENCE   160 AA;  18114 MW;  6E3345141E33C007 CRC64;

  Query Match             39.0%;  Score 336;  DB 1;  Length 160;
  Best Local Similarity   44.4%;  Pred. No. 9.3e-28;
  Matches   59;  Conservative   26;  Mismatches   48;  Indels    0;  Gaps    0;

Qy         27 WPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSKQLKSNINIRFGIILREDIKELFLDLA 86
              |||| || :  ||: |:|||| ||| ||:::||       :||  ::||: ||||:  :
Db         22 WPTHTVCKEENLEIYYKSCDPQQDFAFSIDRCSDVTTHTFDIRAAMVLRQSIKELYAKVD 81

Qy         87 LMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQIYYAGPVNNPEFTIPQGEYQVLLELYT 146
              |:  | :||::|  :|   | |  |||::||| :|| ||:      ||| :| :   |
Db         82 LIINGKTVLSYSETLCGPGLSKLIFCGKKKGEHLYYEGPITLGIKEIPQRDYTITARLTN 141

Qy        147 EKRSTVACANATI 159
              | |:|||||: |:
Db        142 EDRATVACADFTV 154


RESULT 4
LY96_BOVIN
ID   LY96_BOVIN     STANDARD;      PRT;   160 AA.
AC   P58754;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Lymphocyte antigen 96 precursor (MD-2 protein).
GN   LY96 OR MD2.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lymph node;
RA   Guionaud C.T., Dubey C., Zumkehr J.R., Sonstegard T.S., Jungi T.W.;
RT   "Role of bovine TLR2, TLR4 and CD14 in the recognition of bacterial
RT   constituents.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Cooperates with TLR4 in the innate immune response to
CC       bacterial lipopolysaccharide (LPS), and with TLR2 in the response
CC       to cell wall components from Gram-positive and Gram-negative
CC       bacteria. Enhances TLR4-dependent activation of NF-kappa-B. Cells
CC       expressing both MD2 and TLR4, but not TLR4 alone, respond to LPS
CC       (By similarity).
CC   -!- SUBUNIT: Heterogeneous homopolymer formed from homodimers;
CC       disulfide-linked. Belongs to the lipopolysaccharide (LPS)
CC       receptor, a multi-protein complex containing at least CD14, LY96
CC       and TLR4. Binds to the extracellular domains of TLR2 and TLR4 (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Extracellular (By similarity).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF368418; AAL16721.1; -.
DR   InterPro; IPR003172; E1_DerP2_DerF2.
DR   InterPro; IPR007110; Ig-like.
DR   Pfam; PF02221; E1_DerP2_DerF2; 1.
DR   SMART; SM00737; ML; 1.
KW   Immune response; Inflammatory response; Signal; Glycoprotein.
FT   SIGNAL       1     16       POTENTIAL.
FT   CHAIN       17    160       LYMPHOCYTE ANTIGEN 96.
FT   DISULFID    95    105       BY SIMILARITY.
FT   CARBOHYD    26     26       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   114    114       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   150    150       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   160 AA;  18483 MW;  D97760E5FE93B352 CRC64;

  Query Match             14.5%;  Score 125;  DB 1;  Length 160;
  Best Local Similarity   25.6%;  Pred. No. 8e-06;
  Matches   41;  Conservative   29;  Mismatches   70;  Indels   20;  Gaps    7;

Qy         10 LWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQ-DFGFSVEKC--SKQLKSNI 66
              |:: :| |         | |    :|: |   : |  || ::       || |   | | :
Db          6 LFSTLFSSIFTEPREKMW----ICNSSDASLWYNYCDDMKFPISVKVEPCVTIKGTKGKL 61

Qy         67 NIRFGIILREDIKELFLDLALMSQGSSVLNFSYP-----ICEAALPKFSFCGRRKGEQIY 121
              :: :  | | ||::|:|:| :     |: : : |      ||       :|||  ||| :
Db         62 HLYY--IARRDIQKLYLNLHI-----SIKSMTLPMRKEVICREYGGDYSFCGALKGETVN 114

Qy        122 YAGPVNNPEFTIPQGEYQVLLELYT-EKRSTVACANATIM 160
                | :          |:|  ::|  :      : | | ||:
Db        115 TTIPFSFQGIRFSPGQYHCVVEAISGNSEEMLFCLNFTII 154


RESULT 5
LY96_CRIGR
ID   LY96_CRIGR     STANDARD;      PRT;   160 AA.
AC   P58755;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Lymphocyte antigen 96 precursor (MD-2 protein).
GN   LY96 OR MD2.
OS   Cricetulus griseus (Chinese hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Cricetulus.
OX   NCBI_TaxID=10029;
RN   [1]
RP   SEQUENCE FROM N.A., AND VARIANT TYR-95.
RC   TISSUE=Ovarian carcinoma;
RX   MEDLINE=21329172; PubMed=11435474;
RA   Schromm A.B., Lien E., Henneke P., Chow J.C., Yoshimura A., Heine H.,
RA   Latz E., Monks B.G., Schwartz D.A., Miyake K., Golenbock D.T.;
RT   "Molecular genetic analysis of an endotoxin nonresponder mutant cell
RT   line. A point mutation in a conserved region of MD-2 abolishes
RT   endotoxin-induced signaling.";
RL   J. Exp. Med. 194:79-88(2001).
CC   -!- FUNCTION: Cooperates with TLR4 in the innate immune response to
CC       bacterial lipopolysaccharide (LPS), and with TLR2 in the response
CC       to cell wall components from Gram-positive and Gram-negative
CC       bacteria. Enhances TLR4-dependent activation of NF-kappa-B. Cells
CC       expressing both MD2 and TLR4, but not TLR4 alone, respond to LPS
CC       (By similarity).
CC   -!- SUBUNIT: Heterogeneous homopolymer formed from homodimers;
CC       disulfide-linked. Belongs to the lipopolysaccharide (LPS)
CC       receptor, a multi-protein complex containing at least CD14, LY96
CC       and TLR4. Binds to the extracellular domains of TLR2 and TLR4 (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF325501; AAK57984.1; -.
DR   InterPro; IPR003172; E1_DerP2_DerF2.
DR   InterPro; IPR007110; Ig-like.
DR   Pfam; PF02221; E1_DerP2_DerF2; 1.
DR   SMART; SM00737; ML; 1.
KW   Immune response; Inflammatory response; Signal; Glycoprotein;
KW   Polymorphism.
FT   SIGNAL        1     16       POTENTIAL.
FT   CHAIN        17    160       LYMPHOCYTE ANTIGEN 96.
FT   DISULFID     95    105       BY SIMILARITY.
FT   CARBOHYD     26     26       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     77     77       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    101    101       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    150    150       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT      95     95       C -> Y (IN ENDOTOXIN NONRESPONDER).
SQ   SEQUENCE   160 AA;  18357 MW;  0E533B1AA5B46DD6 CRC64;

  Query Match             14.4%;  Score 124;  DB 1;  Length 160;
  Best Local Similarity   25.4%;  Pred. No. 1e-05;
  Matches   33;  Conservative   23;  Mismatches   72;  Indels    2;  Gaps    2;

Qy         33 CSDSGLEVLYQSCDPLQ-DFGFSVEKCSKQLKSNINIRFGIILREDIKELFLDLALMSQG 91
              |:|  |   | || ::  |  : | |     :|  :    |  | |:|:|: :|::
Db         25 CNSSDATFSYSYCDSMKFPFSITAEPCITLKGTNGFLHIKFIPRRDLKKLYFNLSINVNS 84

Qy         92 SSVLNFSYPICEAALPKFSFCGRRKGEQIYYAGPVNNPEFTIPQGEYQVLLE-LYTEKRS 150
                 |      ||      :|||   ||| :   | :      |:|:|: : | :  :
Db         85 IEVPTRKEIICHGYDDNYSFCKALKGETVNTVVPFSFKGILFPKGQYRCVAEAIVGDNEE 144

Qy        151 TVACANATIM 160
               : | | ||:
Db        145 KLFCLNFTII 154


RESULT 6
LY96_MOUSE
ID   LY96_MOUSE     STANDARD;      PRT;   160 AA.
AC   Q9JHF9;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Lymphocyte antigen 96 precursor (MD-2 protein) (ESOP-1).
GN   LY96 OR MD2 OR ESOP1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20352025; PubMed=10891475;
RA   Kato K., Morrison A.M., Nakano T., Tashiro K., Honjo T.;
RT   "ESOP-1, a secreted protein expressed in the hematopoietic, nervous,
RT   and reproductive systems of embryonic and adult mice.";
RL   Blood 96:362-364(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=20191999; PubMed=10725698;
RA   Akashi S., Shimazu R., Ogata H., Nagai Y., Takeda K., Kimoto M.,
RA   Miyake K.;
RT   "Cell surface expression and lipopolysaccaride signaling via the Toll-
RT   like receptor 4-MD-2 complex on mouse peritoneal macrophages.";
RL   J. Immunol. 164:3471-3475(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Embryo;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
CC   -!- FUNCTION: Cooperates with TLR4 in the innate immune response to
CC       bacterial lipopolysaccharide (LPS), and with TLR2 in the response
CC       to cell wall components from Gram-positive and Gram-negative
CC       bacteria. Enhances TLR4-dependent activation of NF-kappa-B. Cells
CC       expressing both MD2 and TLR4, but not TLR4 alone, respond to LPS
CC       (By similarity).
CC   -!- SUBUNIT: Heterogeneous homopolymer formed from homodimers;
CC       disulfide-linked. Belongs to the lipopolysaccharide (LPS)
CC       receptor, a multi-protein complex containing at least CD14, LY96
CC       and TLR4. Binds to the extracellular domains of TLR2 and TLR4 (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- TISSUE SPECIFICITY: Highly expressed in spleen, bone marrow,
CC       thymus, liver, kidney, ovary and decidua. Detected at lower levels
CC       in testis, small intestine and skin.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF168120; AAF89634.1; -.
```

```
DR   EMBL; AB018550; BAA93619.1; -.
DR   EMBL; AK019283; BAB31645.1; -.
DR   MGD; MGI:1341909; Ly96.
DR   InterPro; IPR003172; E1_DerP2_DerF2.
DR   InterPro; IPR007110; Ig-like.
DR   Pfam; PF02221; E1_DerP2_DerF2; 1.
DR   SMART; SM00737; ML; 1.
KW   Immune response; Inflammatory response; Signal; Glycoprotein.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19    160       LYMPHOCYTE ANTIGEN 96.
FT   DISULFID     95    105       BY SIMILARITY.
FT   CARBOHYD     26     26       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    114    114       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    150    150       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   160 AA;  18394 MW;  E224D17E5D5429E2 CRC64;

  Query Match             12.3%;  Score 106;  DB 1;  Length 160;
  Best Local Similarity   26.3%;  Pred. No. 0.00076;
  Matches   36;  Conservative   24;  Mismatches   61;  Indels   16;  Gaps    5;

Qy         33 CSDSGLEVLYQSCDPLQ-DFGFSVEKCSKQLKSNINIRFGIILREDIKELFLDLALMSQG 91
              |: |   : |  || |:     | | | :   :|  :    | | ::| |: :| :
Db         25 CNSSDAIISYSYCDHLKFPISISSEPCIRLRGTNGFVHVEFIPRGNLKYLYFNLFI---- 80

Qy         92 SSVLNFSYP-----ICEAALPKFSFCGRRKGEQIYYAGPVNNPEFTIPQGEYQVLLELY- 145
              || :   |      :|       :|||   ||| :  : | :       |:| |: : |
Db         81 -SVNSIELPKRKEVLCHGHDDDYSFCRALKGETVNTSIPFSFEGILFPKGHYRCVAEAIA 139

Qy        146 --TEKRSTVACANATIM 160
                ||::  : | | ||:
Db        140 GDTEEK--LFCLNFTII 154


RESULT 7
LY96_HUMAN
ID   LY96_HUMAN     STANDARD;      PRT;   160 AA.
AC   Q9Y6Y9;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Lymphocyte antigen 96 precursor (MD-2 protein) (ESOP-1).
GN   LY96 OR MD2 OR ESOP1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Uterus;
RX   MEDLINE=99288036; PubMed=10359581;
RA   Shimazu R., Akashi S., Ogata H., Nagai Y., Fukudome K., Miyake K.,
RA   Kimoto M.;
RT   "MD-2, a molecule that confers lipopolysaccharide responsiveness on
RT   Toll-like receptor 4.";
RL   J. Exp. Med. 189:1777-1782(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20352025; PubMed=10891475;
RA   Kato K., Morrison A.M., Nakano T., Tashiro K., Honjo T.;
RT   "ESOP-1, a secreted protein expressed in the hematopoietic, nervous,
RT   and reproductive systems of embryonic and adult mice.";
RL   Blood 96:362-364(2000).
RN   [3]
RP   SEQUENCE FROM N.A., AND MUTAGENESIS OF CYS-95.
RX   MEDLINE=21329172; PubMed=11435474;
RA   Schromm A.B., Lien E., Henneke P., Chow J.C., Yoshimura A., Heine H.,
RA   Latz E., Monks B.G., Schwartz D.A., Miyake K., Golenbock D.T.;
RT   "Molecular genetic analysis of an endotoxin nonresponder mutant cell
RT   line. A point mutation in a conserved region of MD-2 abolishes
RT   endotoxin-induced signaling.";
RL   J. Exp. Med. 194:79-88(2001).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   INTERACTION WITH TLR2 AND TLR4.
RX   MEDLINE=21103214; PubMed=11160242;
RA   Dziarski R., Wang Q., Miyake K., Kirschning C.J., Gupta D.;
RT   "MD-2 enables Toll-like receptor 2 (TLR2)-mediated responses to
RT   lipopolysaccharide and enhances TLR2-mediated responses to
RT   Gram-positive and Gram-negative bacteria and their cell wall
RT   components.";
RL   J. Immunol. 166:1938-1944(2001).
RN   [6]
RP   DISULFIDE-BONDS, AND CARBOHYDRATE-LINKAGE SITES.
RX   MEDLINE=21477444; PubMed=11593030;
RA   Visintin A., Mazzoni A., Spitzer J.A., Segal D.M.;
RT   "Secreted MD-2 is a large polymeric protein that efficiently confers
RT   lipopolysaccharide sensitivity to Toll-like receptor 4.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:12156-12161(2001).
RN   [7]
RP   INTERCHAIN DISULFIDE BOND.
RX   MEDLINE=22558532; PubMed=12642668;
RA   Mullen G.E.D., Kennedy M.N., Visintin A., Mazzoni A., Leifer C.A.,
RA   Davies D.R., Segal D.M.;
RT   "The role of disulfide bonds in the assembly and function of MD-2.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:3919-3924(2003).
CC   -!- FUNCTION: Cooperates with TLR4 in the innate immune response to
CC       bacterial lipopolysaccharide (LPS), and with TLR2 in the response
CC       to cell wall components from Gram-positive and Gram-negative
CC       bacteria. Enhances TLR4-dependent activation of NF-kappa-B. Cells
CC       expressing both MD2 and TLR4, but not TLR4 alone, respond to LPS.
CC   -!- SUBUNIT: Heterogeneous homopolymer formed from homodimers;
CC       disulfide-linked. Belongs to the lipopolysaccharide (LPS)
CC       receptor, a multi-protein complex containing at least CD14, LY96
CC       and TLR4. Binds to the extracellular domains of TLR2 and TLR4.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- PTM: N-glycosylated; high-mannose.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AB018549; BAA78717.1; -.
DR   EMBL; AF168121; AAF89635.1; -.
DR   EMBL; BC020690; AAH20690.1; -.
DR   Genew; HGNC:17156; LY96.
DR   MIM; 605243; -.
DR   GO; GO:0005886; C:plasma membrane; TAS.
DR   GO; GO:0015026; F:coreceptor activity; TAS.
DR   GO; GO:0006961; P:antibacterial humoral response (sensu Inver. . .; TAS.
```

```
DR   GO; GO:0007166; P:cell surface receptor linked signal transdu. . .; TAS.
DR   GO; GO:0006968; P:cellular defense response; TAS.
DR   InterPro; IPR003172; E1_DerP2_DerF2.
DR   InterPro; IPR007110; Ig-like.
DR   Pfam; PF02221; E1_DerP2_DerF2; 1.
DR   SMART; SM00737; ML; 1.
KW   Immune response; Inflammatory response; Signal; Glycoprotein.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19    160       LYMPHOCYTE ANTIGEN 96.
FT   DISULFID     95    105
FT   CARBOHYD     26     26       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    114    114       N-LINKED (GLCNAC. . .).
FT   MUTAGEN      95     95       C->Y: ABOLISHES LPS-RESPONSE.
SQ   SEQUENCE   160 AA;  18446 MW;  1E83AFF583636D7A CRC64;

  Query Match             11.8%;  Score 101.5;  DB 1;  Length 160;
  Best Local Similarity   24.1%;  Pred. No. 0.0022;
  Matches   34;  Conservative   26;  Mismatches   70;  Indels   11;  Gaps    3;

Qy          4 FTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQ-DFGFSVEKCSKQL 62
              | :|||  : ||       :|   : ||: |   : |  || :|     :|  | :
Db          6 FFSTLF--SSIFT--------EAQKQYWVCNSSDASISYTYCDKMQYPISINVNPCIELK 55

Qy         63 KSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQIYY 122
              |   :    | | |:|:|: :| :     ::      ||  :   :|||   ||| :
Db         56 GSKGLLHIFYIPRRDLKQLYFNLYIITVNTMNLPKRKEVICRGSDDDYSFCRALKGETVNT 115

Qy        123 AGPVNNPEFTIPQGEYQVLLE 143
                 :       :|:|: ::|
Db        116 TISFSFKGIKFSKGKYKCVVE 136


RESULT 8
TOP1_VIBCH
ID   TOP1_VIBCH     STANDARD;      PRT;   876 AA.
AC   Q9KRB2;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   DNA topoisomerase I (EC 5.99.1.2) (Omega-protein) (Relaxing enzyme)
DE   (Untwisting enzyme) (Swivelase).
GN   TOPA OR VC1730.
OS   Vibrio cholerae.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC   Vibrionaceae; Vibrio.
OX   NCBI_TaxID=666;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=El Tor N16961 / Serotype O1;
RX   MEDLINE=20406833; PubMed=10952301;
RA   Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA   Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA   Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA   Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA   McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA   Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA   Fraser C.M.;
RT   "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT   cholerae.";
RL   Nature 406:477-483(2000).
CC   -!- FUNCTION: The reaction catalyzed by topoisomerases leads to the
CC       conversion of one topological isomer of DNA to another.
CC   -!- CATALYTIC ACTIVITY: ATP-independent breakage of single-stranded
CC       DNA, followed by passage and rejoining.
CC   -!- SUBUNIT: Monomer (By similarity).
CC   -!- MISCELLANEOUS: When a topoisomerase transiently breaks a DNA
CC       backbone bond, it simultaneously forms a protein-DNA link, in
CC       which a tyrosyl oxygen in the enzyme is joined to a DNA phosphorus
CC       at one end of the enzyme-severed DNA strand.
CC   -!- SIMILARITY: Belongs to the prokaryotic type I/III topoisomerase
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE004251; AAF94880.1; -.
DR   PIR; B82163; B82163.
DR   HSSP; P06612; 1ECL.
DR   TIGR; VC1730; -.
DR   InterPro; IPR005733; DNA_topI_bact.
DR   InterPro; IPR000380; DNA_tpisomrase.
DR   InterPro; IPR003601; DNAtopI_ATP_bind.
DR   InterPro; IPR003602; DNAtopI_DNA_bind.
DR   InterPro; IPR006171; Toprim_dom.
DR   InterPro; IPR006154; Toprim_sub.
DR   Pfam; PF01131; Topoisom_bac; 1.
DR   Pfam; PF01751; Toprim; 1.
DR   Pfam; PF01396; zf-C4_Topoisom; 2.
DR   PRINTS; PR00417; PRTPISMRASEI.
DR   SMART; SM00437; TOP1Ac; 1.
DR   SMART; SM00436; TOP1Bc; 1.
DR   SMART; SM00493; TOPRIM; 1.
DR   TIGRFAMs; TIGR01051; topA_bact; 1.
DR   PROSITE; PS00396; TOPOISOMERASE_I_PROK; 1.
KW   Isomerase; Topoisomerase; DNA-binding; Zinc-finger; Metal-binding;
KW   Repeat; Complete proteome.
FT   ZN_FING     668    695       C4-TYPE 1.
FT   ZN_FING     717    742       C4-TYPE 2.
FT   ACT_SITE    327    327       DNA CLEAVAGE (BY SIMILARITY).
SQ   SEQUENCE   876 AA;  97909 MW;  8A5BC9148CA1B753 CRC64;

  Query Match              9.0%;  Score 78;  DB 1;  Length 876;
  Best Local Similarity   26.1%;  Pred. No. 4.9;
  Matches   37;  Conservative   23;  Mismatches   40;  Indels   42;  Gaps   10;

Qy          8 LFLWTLIFPSCSGGGGGKAWPTHVV--CSDSGLEVLYQSCDPLQDFGFSVEKCSKQLKSN 65
              : :  :: |:||   | :   | |   ||  ||                 |:|    |:
Db        599 IVMTNILCPTCSRPMGIRTASTGVFLGCSGYGLP-------------PKERC----KTT 640

Qy         66 INI--RFGII--LREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQIY 121
              ||:    |:|  | ||::      || ::          |||| |:  :    :||   ::
Db        641 INLGDEEGVINVLEEDVE----TAALRAKKRC------PICETAMDAYLIDDKRK---LH 687

Qy        122 YAGPVNNPE---FTIPQGEYQV 140
                |  |||    | : :||::|
Db        688 VCG--NNPNCEGFIVEEGEFKV 707


RESULT 9
CYSC_VIBPA
ID   CYSC_VIBPA     STANDARD;      PRT;   205 AA.
AC   Q87SX6;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Adenylylsulfate kinase (EC 2.7.1.25) (APS kinase) (Adenosine-
DE   5'phosphosulfate kinase) (ATP adenosine-5'-phosphosulfate 3'-
DE   phosphotransferase).
GN   CYSC OR VP0296.
OS   Vibrio parahaemolyticus.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC   Vibrionaceae; Vibrio.
OX   NCBI_TaxID=670;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RIMD 2210633 / Serotype O3:K6;
RX   MEDLINE=22508454; PubMed=12620739;
RA   Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA   Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA   Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT   "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
```

```
RT   distinct from that of V. cholerae.";
RL   Lancet 361:743-749(2003).
CC   -!- FUNCTION: Catalyzes the synthesis of activated sulfate.
CC   -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-
CC       phosphoadenylylsulfate.
CC   -!- PATHWAY: Sulfate activation; cysteine biosynthesis reductive
CC       branch; second step.
CC   -!- SIMILARITY: Belongs to the APS kinase family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AP005073; BAC58559.1; -.
DR   HAMAP; MF_00065; -; 1.
DR   InterPro; IPR002891; APS_kinase.
DR   Pfam; PF01583; APS_kinase; 1.
KW   Transferase; Kinase; Cysteine biosynthesis; ATP-binding;
KW   Phosphorylation; Complete proteome.
FT   NP_BIND      39     46       ATP (BY SIMILARITY).
FT   ACT_SITE    113    113       PHOSPHOSERINE INTERMEDIATE (BY
FT                                SIMILARITY).
SQ   SEQUENCE   205 AA;  22894 MW;  03696F141F48199F CRC64;

  Query Match             8.9%;  Score 77;  DB 1;  Length 205;
  Best Local Similarity   26.3%;  Pred. No. 1.1;
  Matches   50;  Conservative   20;  Mismatches   56;  Indels   64;  Gaps   14;

Qy          2 KGFTATL------FLWTLIFPSCSGGGGGK-----------AWPTHVVCSDSGLEVLYQ 43
              | | | |       ||   |   || |                : |:::  |:   | :
Db         21 KQFRAELKKQKPAVLW---FTGLSGAGKSTVAGALENRLAELGYHTYLLDGDN---VRHG 74

Qy         44 SCDPLQDFGFSVEKCSKQLKSNINIRFG-----------IILREDI------KELFLDLA 86
                |    | |||     :  : ||  | |           |:|   |      ::|  |
Db         75 LC---SDLGFS----EQDRRENIR-RIGELAKLMADAGLIVLSAFISPHRAERQLVRD-- 124

Qy         87 LMSQGSSV---LNFSYPICEAALPKFSFCGRRKGE-------QIYYAGPVNNPEFTIPQG 136
              |: :|  :   :| |  :||   ||  :   | ||           |  |: ||  :| |
Db        125 LLPEGEFIEVFVNASLEVCEGRDPKGLYKKARAGEIPNFTGIDSEYQAPI-NPEIDLPAG 183

Qy        137 EYQV--LLEL 144
              |  |  |:||
Db        184 EKSVEELVEL 193


RESULT 10
CYSC_VIBVU
ID   CYSC_VIBVU     STANDARD;      PRT;   207 AA.
AC   Q8DE75;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Adenylylsulfate kinase (EC 2.7.1.25) (APS kinase) (Adenosine-
DE   5'phosphosulfate kinase) (ATP adenosine-5'-phosphosulfate 3'-
DE   phosphotransferase).
GN   CYSC OR VV10723.
OS   Vibrio vulnificus.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC   Vibrionaceae; Vibrio.
OX   NCBI_TaxID=672;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CMCP6;
RA   Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA   Choy H.E.;
RT   "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL   Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Catalyzes the synthesis of activated sulfate.
CC   -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-
CC       phosphoadenylylsulfate.
CC   -!- PATHWAY: Sulfate activation; cysteine biosynthesis reductive
CC       branch; second step.
CC   -!- SIMILARITY: Belongs to the APS kinase family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE016799; AAO09232.1; -.
DR   HAMAP; MF_00065; -; 1.
DR   InterPro; IPR002891; APS_kinase.
DR   Pfam; PF01583; APS_kinase; 1.
DR   ProDom; PD002350; APS_kinase; 1.
DR   TIGRFAMs; TIGR00455; apsK; 1.
KW   Transferase; Kinase; Cysteine biosynthesis; ATP-binding;
KW   Phosphorylation; Complete proteome.
FT   NP_BIND      39     46       ATP (BY SIMILARITY).
FT   ACT_SITE    113    113       PHOSPHOSERINE INTERMEDIATE
FT                                (BY SIMILARITY).
SQ   SEQUENCE   207 AA;  22886 MW;  225B1E4711939951B CRC64;

  Query Match             8.9%;  Score 77;  DB 1;  Length 207;
  Best Local Similarity   25.0%;  Pred. No. 1.1;
  Matches   45;  Conservative   26;  Mismatches   51;  Indels   58;  Gaps   13;

Qy         10 LWTLIFPSCSGGG-----------GGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEK 57
              ||   |  || |               : : |:::  |:   | :  |    | |||
Db         35 LW---FTGLSGAGKSTVAGALENRLAEQGFHTYLLDGDN---VRHGLC---SDLGFS--- 82

Qy         58 CSKQLKSNINIRFG-----------IILREDI------KELFLDLALMSQGSSV---LNF 97
              ::  : ||  | |             |:|   |      ::|  |  |: :|  :   :|
Db         83 -TQDRRENIR-RIGELAKLMADAGLIVLTAFISPHRAERQLVRD--LLPEGEFIEVFVNT 138

Qy         98 SYPICEAALPKFSFCGRRKGEQIYYAG-------PVNNPEFTIPQGE--YQVLLELYTEK 148
              |  :||   ||  :   | ||    : |       |: |||  :| ||   : |::|  |:
Db        139 SLEVCEQRDPKGLYKKARAGEIANFTGIDSEYEVPL-NPEIDLPAGEKGIEALVDLLVEQ 197


RESULT 11
VGLD_HSVEA
ID   VGLD_HSVEA     STANDARD;      PRT;   402 AA.
AC   P24872;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Glycoprotein D precursor (Glycoprotein 17/18).
GN   GD OR GP17/18 OR 72.
OS   Equine herpesvirus type 1 (strain AB1) (EHV-1).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Varicellovirus.
OX   NCBI_TaxID=10328;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92268882; PubMed=1316942;
RA   Elton D.M., Halliburton I.W., Killington R.A., Meredith D.M.,
RA   Bonass W.A.;
RT   "Identification of the equine herpesvirus type 1 glycoprotein 17/18
RT   as a homologue of herpes simplex virus glycoprotein D.";
RL   J. Gen. Virol. 73:1227-1233(1992).
RN   [2]
RP   SEQUENCE OF 242-402 FROM N.A.
RX   MEDLINE=91276272; PubMed=1647359;
RA   Elton D.M., Halliburton I.W., Killington R.A., Meredith D.M.,
RA   Bonass W.A.;
RT   "Sequence analysis of the 4.7-kb BamHI-EcoRI fragment of the equine
RT   herpesvirus type-1 short unique region.";
RL   Gene 101:203-208(1991).
CC   -!- SIMILARITY: Belongs to the herpesviruses glycoprotein D family.
```

```
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M60946; AAA46087.1; -.
DR   EMBL; M36299; AAA66546.1; -.
DR   InterPro; IPR002896; Herpes_glycop_D.
DR   InterPro; IPR007110; Ig-like.
DR   Pfam; PF01537; Herpes_glycop_D; 1.
KW   Glycoprotein; Transmembrane; Signal.
FT   SIGNAL        1     30       POTENTIAL.
FT   CHAIN        31    402       GLYCOPROTEIN D.
FT   DOMAIN       31    355       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    356    372       POTENTIAL.
FT   DOMAIN      373    402       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     53     53       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     61     61       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    297    297       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    346    346       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   402 AA;  45211 MW;  78A0593232D0238C CRC64;

  Query Match             8.9%;  Score 76.5;  DB 1;  Length 402;
  Best Local Similarity   26.0%;  Pred. No. 2.7;
  Matches   44;  Conservative   16;  Mismatches   62;  Indels   47;  Gaps    8;

Qy          2 KGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSKQ 61
              | ::| |  |  | |:|       | | |       :| |  |:|   |    |
Db        111 KTYSARL-TWFKIMPTC-------ATPIH--------DVSYMKCNPKLSFAMCDE----- 149

Qy         62 LKSNINIRFGII--LREDIKELFLDLALMSQGSSVL-------------NFSYPICEAA 105
               :|:|  :  :|    |   || | ||  :  :| |             :||  |
Db        150 -RSDILWQASLITMAAETDDELGLVLAAPAHSASGLYRRVIEIDGRQIYTDFSVTIPSER 208

Qy        106 LP-----KFSFCGRRKGEQIYYAGPVNN----PEFTIPQGEYQVLLELY 145
              |     |    | |  : |  | |       ||  ||||:  .|: :
Db        209 CPIAFEQNFGNPDRCKTPEQYSRGEVFTRRFLGEFNFPQGEHMTWLKFW 257


RESULT 12
VGLD_HSVEB
ID   VGLD_HSVEB     STANDARD;      PRT;   452 AA.
AC   P24379;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Glycoprotein D precursor (Glycoprotein 17/18).
GN   GD OR GP17/18 OR 72.
OS   Equine herpesvirus type 1 (strain Ab4p) (EHV-1), and
OS   Equine herpesvirus type 1 (strain Kentucky D) (EHV-1).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Varicellovirus.
OX   NCBI_TaxID=31520, 10330;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB4P;
RX   MEDLINE=92295566; PubMed=1318606;
RA   Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
RT   "The DNA sequence of equine herpesvirus-1.";
RL   Virology 189:304-316(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Kentucky D;
RX   MEDLINE=91108393; PubMed=2177089;
RA   Audonnet J.-C., Winslow J., Allen G., Paoletti E.;
RT   "Equine herpesvirus type 1 unique short fragment encodes
RT   glycoproteins with homology to herpes simplex virus type 1 gD, gI and
RT   gE.";
RL   J. Gen. Virol. 71:2969-2978(1990).
CC   -!- SIMILARITY: Belongs to the herpesviruses glycoprotein D family.
CC   -!- CAUTION: It is uncertain whether Met-1 or Met-51 is the initiator.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M86664; AAB02507.1; -.
DR   PIR; I36802; VGBEG3.
DR   InterPro; IPR002896; Herpes_glycop_D.
DR   InterPro; IPR007110; Ig-like.
DR   Pfam; PF01537; Herpes_glycop_D; 1.
KW   Glycoprotein; Signal; Transmembrane.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    452       GLYCOPROTEIN D.
FT   DOMAIN       20    405       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    406    422       POTENTIAL.
FT   DOMAIN      423    452       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    103    103       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    111    111       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    347    347       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    396    396       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   452 AA;  51099 MW;  CF51E914F7F2E9DC CRC64;

  Query Match             8.9%;  Score 76.5;  DB 1;  Length 452;
  Best Local Similarity   26.0%;  Pred. No. 3.1;
  Matches   44;  Conservative   16;  Mismatches   62;  Indels   47;  Gaps    8;

Qy          2 KGFTATLFLWTLIFPSCSGGGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSKQ 61
              | ::| |  |  | |:|       | | |       :| |  |:|   |    |
Db        161 KTYSARL-TWFKIMPTC-------ATPIH--------DVSYMKCNPKLSFAMCDE----- 199

Qy         62 LKSNINIRFGII--LREDIKELFLDLALMSQGSSVL-------------NFSYPICEAA 105
               :|:|  :  :|    |   || | ||  :  :| |             :||  |
Db        200 -RSDILWQASLITMAAETDDELGLVLAAPAHSASGLYRRVIEIDGRRIYTDFSVTIPSER 258

Qy        106 LP-----KFSFCGRRKGEQIYYAGPVNN----PEFTIPQGEYQVLLELY 145
              |     |    | |  : |  | |       ||  ||||:   |: :
Db        259 CPIAFEQNFGNPDRCKTPEQYSRGEVFTRRFLGEFNFPQGEHMTWLKFW 307


RESULT 13
MURG_CAMJE
ID   MURG_CAMJE     STANDARD;      PRT;   342 AA.
AC   Q9PNQ2;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide)
DE   pyrophosphoryl-undecaprenol N-acetylglucosamine transferase
DE   (EC 2.4.1.227) (Undecaprenyl-PP-MurNAc-pentapeptide-UDPGlcNAc GlcNAc
DE   transferase).
GN   MURG OR CJ1039.
OS   Campylobacter jejuni.
OC   Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC   Campylobacteraceae; Campylobacter.
OX   NCBI_TaxID=197;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCTC 11168;
RX   MEDLINE=20150912; PubMed=10688204;
RA   Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA   Basham D., Chillingworth T., Davies R.M., Feltwell T., Holroyd S.,
RA   Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA   Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA   Whitehead S., Barrell B.G.;
RT   "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT   reveals hypervariable sequences.";
RL   Nature 403:665-668(2000).
```

```
CC   -!- FUNCTION: Cell wall formation. Catalyzes the transfer of a GlcNAc
CC       subunit on undecaprenyl-pyrophosphoryl-MurNAc-pentapeptide (lipid
CC       intermediate I) to form undecaprenyl-pyrophosphoryl-MurNAc-
CC       (pentapeptide)GlcNAc (lipid intermediate II) (By similarity).
CC   -!- CATALYTIC ACTIVITY: UDP-N-acetylglucosamine + Mur2Ac(oyl-L-Ala-
CC       gamma-D-Glu-L-Lys-D-Ala-D-Ala)-diphosphoundecaprenol = UDP +
CC       GlcNAc-(1->4)-Mur2Ac(oyl-L-Ala-gamma-D-Glu-L-Lys-D-Ala-D-Ala)-
CC       diphosphoundecaprenol.
CC   -!- PATHWAY: Peptidoglycan biosynthesis; last step.
CC   -!- SUBCELLULAR LOCATION: Inner membrane-associated (By similarity).
CC   -!- SIMILARITY: Belongs to the glycosyltransferase family 28. MurG
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bicinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AL139077; CAB73295.1; -.
DR   PIR; D81306; D81306.
DR   HAMAP; MF_00033; -; 1.
DR   InterPro; IPR007235; Glyco_tran_28_C.
DR   InterPro; IPR004276; Glyco_trans_28.
DR   InterPro; IPR006009; MurG.
DR   Pfam; PF04101; Glyco_tran_28_C; 1.
DR   Pfam; PF03033; Glyco_transf_28; 1.
DR   TIGRFAMs; TIGR01133; murG; 1.
KW   Transferase; Glycosyltransferase; Cell division; Cell wall; Membrane;
KW   Inner membrane; Peptidoglycan synthesis; Complete proteome.
SQ   SEQUENCE   342 AA;  38683 MW;  F70131387068E9E7 CRC64;

  Query Match             8.8%;  Score 75.5;  DB 1;  Length 342;
  Best Local Similarity   25.7%;  Pred. No. 2.8;
  Matches   37;  Conservative   22;  Mismatches   34;  Indels   51;  Gaps   11;

Qy         35 DSGLEVLYQSCDPLQDFGFSVEKCSKQLKSNINIR-----FGIILREDIKELFLDLALMS 89
              :  :::::| |    ||    ||| |  :| :||:     | : | | :|    |||:
Db        196 EQNIKIIHQ-CGK-NDF----EKCKKHYQS-LNIQADIFDFSLNLEEKMKN--ADLAISR 246

Qy         90 QGSSVLNFSYPICEAALP-----------------KF----SFCGRRKGEQIYYAGPVN 127
                |:| |   : :|   ||                 ||    : |      ||:    :|
Db        247 AGASTL---FELCANTLPTIFIPYPYAAKNHQYFNAKFLQDQALC------QIFMQNSIN 297

Qy        128 NPEFTIPQGEYQVLLELYTEKRST 151
                ||      :: :|:|  |  ||
Db        298 LDEF------FKSILKLNLENIST 315


RESULT 14
RPOB_CLOAB
ID   RPOB_CLOAB     STANDARD;      PRT;  1241 AA.
AC   Q97EG9;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   DNA-directed RNA polymerase beta chain (EC 2.7.7.6) (Transcriptase
DE   beta chain) (RNA polymerase beta subunit).
GN   RPOB OR CAC3143.
OS   Clostridium acetobutylicum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1488;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX   MEDLINE=21359325; PubMed=11466286;
RA   Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA   Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA   Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA   Bennett G.N., Koonin E.V., Smith D.R.;
RT   "Genome sequence and comparative analysis of the solvent-producing
RT   bacterium Clostridium acetobutylicum.";
RL   J. Bacteriol. 183:4823-4838(2001).
CC   -!- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
CC       of DNA into RNA using the four ribonucleoside triphosphates as
CC       substrates.
CC   -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC       (RNA)(N).
CC   -!- SUBUNIT: The enzyme consists of the sigma chain and the core
CC       enzyme which is composed of 2 alpha chains, 1 beta chain, and 1
CC       beta' chain.
CC   -!- SIMILARITY: Belongs to the RNA polymerase beta chain family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE007809; AAK81081.1; -.
DR   PIR; F97286; F97286.
DR   InterPro; IPR007121; RNA_pol_B.
DR   InterPro; IPR007644; RNA_pol_Rpb2_1.
DR   InterPro; IPR007642; RNA_pol_Rpb2_2.
DR   InterPro; IPR007645; RNA_pol_Rpb2_3.
DR   InterPro; IPR007120; RNA_pol_Rpb2_6.
DR   InterPro; IPR007641; RNA_pol_Rpb2_7.
DR   Pfam; PF04563; RNA_pol_Rpb2_1; 1.
DR   Pfam; PF04561; RNA_pol_Rpb2_2; 2.
DR   Pfam; PF04565; RNA_pol_Rpb2_3; 1.
DR   Pfam; PF00562; RNA_pol_Rpb2_6; 1.
DR   Pfam; PF04560; RNA_pol_Rpb2_7; 1.
DR   PROSITE; PS01166; RNA_POL_BETA; 1.
KW   Transferase; Transcription; DNA-directed RNA polymerase;
KW   Complete proteome.
SQ   SEQUENCE   1241 AA;  139248 MW;  DEA50BA38EC4CB78 CRC64;

  Query Match             8.8%;  Score 75.5;  DB 1;  Length 1241;
  Best Local Similarity   19.3%;  Pred. No. 14;
  Matches   29;  Conservative   42;  Mismatches   38;  Indels   41;  Gaps    9;

Qy         35 DSGLEVLYQSCDPLQDF-----------------GFSVEKCSKQ---LKSNINIRFGII 73
              |:||: ::   :|:||:                 :|||:| ::     : : ::  ::
Db         40 DNGLQEIFDDINPIQDYTGNLILEFIGYKLDMDNIKYSVEECKERDTTYAAPLKVKVRLL 99

Qy         74 LRE--DIK--ELFL-DLALMS-QGSSVLNFSYPICEAALPK-----FSFCGRRKGEQIYY 122
               :|  ::|  |:|: |  || |: ::| :  :  : | :     : :   : |::::
Db        100 NKETGEVKEQEVFMGDFPLMTEQGTFIINGAERVIVSQLVRSPGAYYDYIVDKNGKKLFS 159

Qy        123 AGPVNNPEFTIPQGEYQVLLELYTEKRSTV 152
              |  : |      :| :   ||  |: | :
Db        160 ATVIPN------RGAW---LEYETDSNSVI 180


RESULT 15
UAP1_HUMAN
ID   UAP1_HUMAN     STANDARD;      PRT;   522 AA.
AC   Q16222;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   UDP-N-acetylhexosamine pyrophosphorylase (Antigen X) (AGX) (Sperm-
DE   associated antigen 2) [Includes: UDP-N-acetylgalactosamine
DE   pyrophosphorylase (EC 2.7.7.-) (AGX-1); UDP-N-acetylglucosamine
DE   pyrophosphorylase (EC 2.7.7.23) (AGX-2)].
GN   UAP1 OR SPAG2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
```

```
RP   SEQUENCE FROM N.A. (ISOFORMS AGX1 AND AGX2).
RC   TISSUE=Testis;
RX   MEDLINE=94297079; PubMed=8025165;
RA   Diekman A.B., Goldberg E.;
RT   "Characterization of a human antigen with sera from infertile
RT   patients.";
RL   Biol. Reprod. 50:1087-1093(1994).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM AGX1).
RC   TISSUE=Testis;
RX   MEDLINE=98269105; PubMed=9603950;
RA   Mio T., Yabe T., Arisawa M., Yamada-Okabe H.;
RT   "The eukaryotic UDP-N-acetylglucosamine pyrophosphorylases: gene
RT   cloning, protein expression, and catalytic mechanism.";
RL   J. Biol. Chem. 273:14392-14397(1998).
RN   [3]
RP   SEQUENCE FROM N.A. (AGX1/AGX2), PARTIAL SEQUENCE, AND
RP   CHARACTERIZATION.
RC   TISSUE=Breast cancer;
RX   MEDLINE=98438464; PubMed=9765219;
RA   Wang-Gillam A., Pastuszak I., Elbein A.D.;
RT   "A 17-amino acid insert changes UDP-N-acetylhexosamine
RT   pyrophosphorylase specificity from UDP-GalNAc to UDP-GlcNAc.";
RL   J. Biol. Chem. 273:27055-27057(1998).
CC   -!- FUNCTION: AGX1 converts GalNaC-1-P into UDP-GalNac in the presence
CC       of UTP. AGX2 catalyzes the synthesis of GlcNAc-1-P into UDP-GlcNAc
CC       in the presence of UTP.
CC   -!- CATALYTIC ACTIVITY: UTP + N-acetyl-alpha-D-glucosamine 1-phosphate
CC       = diphosphate + UDP-N-acetyl-D-glucosamine.
CC   -!- PATHWAY: UDP-GlcNAc biosynthesis from Fru-6-P; fourth (last) step.
CC   -!- SUBUNIT: Homodimer.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC. IN SPERMATOZOA, LOCALIZED TO
CC       THE PRINCIPAL PIECE OF THE TAIL, THE NECK REGION OF THE HEAD AND
CC       TO A LESSER EXTENT, THE MIDPIECE OF THE TAIL.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=AGX2; Synonyms=AGX-2;
CC         IsoId=Q16222-1; Sequence=Displayed;
CC       Name=AGX1; Synonyms=AGX-1;
CC         IsoId=Q16222-2; Sequence=VSP_004483;
CC   -!- TISSUE SPECIFICITY: Widely expressed. AGX1 is more abundant in
CC       testis than AGX2, while AGX2 is more abundant than AGX1 in somatic
CC       tissue. Expressed at low level in placenta, muscle and liver.
CC   -!- DISEASE: ANTIGEN IMPLICATED IN ANTIBODY-MEDIATED MALE INFERTILITY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; S73498; AAB31210.2; -.
DR   EMBL; AB011004; BAA31202.1; -.
DR   PDB; 1JV1; 28-AUG-02.
DR   PDB; 1JV3; 28-AUG-02.
DR   PDB; 1JVD; 29-AUG-02.
DR   PDB; 1JVG; 30-AUG-02.
DR   Genew; HGNC:12457; UAP1.
DR   MIM; 602862; -.
DR   GO; GO:0003977; F:UDP-N-acetylglucosamine diphosphorylase act. . .; TAS.
DR   GO; GO:0006048; P:UDP-N-acetylglucosamine biosynthesis; TAS.
DR   InterPro; IPR002618; UDPGP.
DR   Pfam; PF01704; UDPGP; 1.
KW   Transferase; Nucleotidyltransferase; Antigen; Alternative splicing;
KW   Polymorphism; 3D-structure.
FT   SITE        111    111       BINDING SITE FOR HEXNAC-1-P
FT                                 (BY SIMILARITY).
FT   ACT_SITE    115    115       POTENTIAL.
FT   ACT_SITE    122    122       POTENTIAL.
FT   VARSPLIC    454    470       Missing (in isoform AGX1).
FT                                 /FTId=VSP_004483.
FT   VARIANT     418    418       P -> H (in dbSNP:1128539).
FT                                 /FTId=VAR_014935.
FT   CONFLICT    454    454       MISSING (IN REF. 1).
SQ   SEQUENCE   522 AA;  58840 MW;  C823A9AD8659A135 CRC64;

  Query Match             8.7%;  Score 75;  DB 1;  Length 522;
  Best Local Similarity   23.7%;  Pred. No. 5.4;
  Matches   28;  Conservative   19;  Mismatches   35;  Indels   36;  Gaps    6;

Qy         27 WPTHVVCSDSGLEVLYQSCDPLQDFGFSVEK---CSKQL--KSNINIRFGIILRED---- 77
              |  || | |:   :| :  || :  || ::|    |  ::  |:|      |:: | |
Db        245 WSIHVYCVDN---ILVKVADP-RFIGFCIQKGADCGAKVVEKTNPTEPVGVVCRVDGVYQ 300

Qy         78 -IKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCGRRKGEQIYYAGPVNNPEFTIP 134
               ::   : ||   : ||                          |  :: || : |  ||:|
Db        301 VVEYSEISLATAQKRSS--------------------DGRLLFNAGNIANHFFTVP 336


Search completed: April 29, 2004, 15:24:51
Job time : 18 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: April 29, 2004, 15:21:06 ; Search time 41 Seconds
(without alignments)
1246.682 Million cell updates/sec

Title: US-09-899-917-2
Perfect score: 862
Sequence: 1 MKGPTATLFLWTLIFPSCSG.....ELYTEKRSTVACANATIMCS 162

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	115	13.3	160	6	Q8MIQ1	Q8miq1 oryctolagus
2	99	11.5	1137	10	O82196	O82196 arabidopsis
3	83	9.6	204	16	Q7TUM5	Q7tum5 prochloroco
4	83	9.6	1440	10	Q9LV30	Q9lv30 arabidopsis
5	81	9.4	406	17	O58069	O58069 pyrococcus
6	79.5	9.2	484	10	Q9SIN9	Q9sin9 arabidopsis
7	79	9.2	2786	5	Q9VSU2	Q9vsu2 drosophila
8	78	9.0	1133	3	Q12139	Q12139 saccharomyc
9	78	9.0	2409	5	Q960G6	Q960g6 drosophila
10	77.5	9.0	394	17	Q9UY77	Q9uy77 pyrococcus
11	77.5	9.0	428	10	O23839	O23839 brassica ol
12	77	8.9	246	16	Q8ZBY9	Q8zby9 yersinia pe
13	77	8.9	396	5	O76660	O76660 toxoplasma
14	77	8.9	480	13	Q7T2R7	Q7t2r7 fugu rubrip
15	77	8.9	739	10	Q9M9G4	Q9m9g4 arabidopsis
16	77	8.9	751	10	Q8RVQ6	Q8rvq6 arabidopsis
17	77	8.9	752	10	Q94AI6	Q94ai6 arabidopsis
18	76.5	8.9	264	5	Q20250	Q20250 caenorhabdi
19	76.5	8.9	402	12	Q04245	Q04245 equine herp
20	76.5	8.9	427	10	Q8S1D7	Q8s1d7 oryza sativ
21	76.5	8.9	436	5	Q23653	Q23653 caenorhabdi
22	76.5	8.9	471	10	Q9MA46	Q9ma46 arabidopsis
23	76.5	8.9	1177	16	Q81VT8	Q81vt8 bacillus an
24	76.5	8.9	1177	16	Q81J48	Q81j48 bacillus ce
25	76	8.8	303	10	Q8W4Y5	Q8w4y5 lycopersico
26	76	8.8	670	10	Q8GX23	Q8gx23 arabidopsis
27	75.5	8.8	534	10	Q8RWH6	Q8rwh6 arabidopsis
28	75.5	8.8	1215	10	P93750	P93750 arabidopsis
29	75	8.7	377	6	O46493	O46493 bos taurus
30	75	8.7	505	4	Q96GM2	Q96gm2 homo sapien
31	75	8.7	699	10	Q9XI96	Q9xi96 arabidopsis
32	74.5	8.6	448	5	Q21761	Q21761 caenorhabdi
33	74.5	8.6	691	10	Q7XNC0	Q7xnc0 oryza sativ
34	74	8.6	374	6	Q9TTW0	Q9ttw0 bos taurus
35	74	8.6	680	5	Q8WSN8	Q8wsn8 caenorhabdi
36	73.5	8.5	340	11	Q8R103	Q8r103 mus musculu
37	73.5	8.5	481	16	Q97D30	Q97d30 clostridium
38	73.5	8.5	601	11	Q8CFC9	Q8cfc9 rattus norv
39	73.5	8.5	995	11	Q8K2R3	Q8k2r3 mus musculu
40	73	8.5	1106	16	Q92H19	Q92h19 rickettsia
41	73	8.5	2382	5	Q9BI19	Q9bi19 drosophila
42	72.5	8.4	220	16	Q8D173	Q8d173 yersinia pe
43	72.5	8.4	314	16	Q7VHI5	Q7vhi5 helicobacte
44	72.5	8.4	553	16	Q8EDP6	Q8edp6 shewanella
45	72.5	8.4	804	6	Q29113	Q29113 sus scrofa

ALIGNMENTS

RESULT 1
Q8MIQ1
ID Q8MIQ1 PRELIMINARY; PRT; 160 AA.
AC Q8MIQ1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MD-2.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Kajikawa O., Frevert C.W., Martin T.R.;
RT "Molecular cloning of rabbit MD-2.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY101395; AAM50061.1; -.
DR InterPro; IPR003172; E1_DerP2_DerF2.
DR InterPro; IPR007110; Ig-like.
DR Pfam; PF02221; E1_DerP2_DerF2; 1.
DR SMART; SM00737; ML; 1.
SQ SEQUENCE 160 AA; 18186 MW; A2C1E0B46E762F43 CRC64;

Query Match 13.3%; Score 115; DB 6; Length 160;
Best Local Similarity 25.4%; Pred. No. 0.00027;
Matches 35; Conservative 26; Mismatches 65; Indels 12; Gaps 4;

```
Qy         30 HVVCSDSGLEVLYQSCDPLQ-DFGFSVEKCSKQLKSNINIRFGIILREDIKELFLDLALM 88
              | ||  :   | |  ||  ::     :|| |     :   :    | | |: :|:|:| :
Db         22 HWVCKSADATVSYTYCDNMEIPISINVEPCITLKGTQGLLHIFYIPRRDMNQLYLNLYI- 80

Qy         89 SQGSSVLNFSYP-----ICEAALPKFSFCGRRKGEQIYYAGPVNNPEFTIPQGEYQVLLE 143
                 || :   |     ||: :   :|||   ||| :    | :     : :|:|: ::|
Db         81 ----SVNSMDLPKRKEIICKGSDDVYSFCRALKGETVNTTVPFSFKGIRLSKGQYRCVVE 136

Qy        144 -LYTEKRSTVACANATIM 160
               :       : | | ||:
Db        137 AIAGSAEEMIFCLNFTII 154
```

```
RESULT 2
O82196
ID   O82196      PRELIMINARY;      PRT;  1137 AA.
AC   O82196;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Copia-like retroelement pol polyprotein.
GN   AT2G19840.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=20083487; PubMed=10617197;
RA   Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA   Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA   Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA   Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA   Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA   Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA   Salzberg S.L., Fraser C.M., Venter J.C.;
RT   "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:761-768(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Lin X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC005169; AAC62132.1; -.
DR   PIR; G84581; G84581.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0006310; P:DNA recombination; IEA.
DR   InterPro; IPR001584; Rve.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF00665; rve; 1.
DR   Pfam; PF00098; zf-CCHC; 1.
DR   PRINTS; PR00939; C2HCZNFINGER.
DR   SMART; SM00343; ZnF_C2HC; 1.
DR   PROSITE; PS50158; ZF_CCHC; 1.
KW   Polyprotein.
SQ   SEQUENCE   1137 AA;  129775 MW;  7265DF09C6E1025C CRC64;

  Query Match             11.5%;  Score 99;  DB 10;  Length 1137;
  Best Local Similarity   26.3%;  Pred. No. 0.15;
  Matches   36;  Conservative   23;  Mismatches   42;  Indels   36;  Gaps    7;

Qy          10 LWTLIFPSCSG----------GGGGKAWPTHVVCSDSGLEVLYQSCDPLQDFGFSV--- 55
               |||   |: ||          || |     :  |: | |    |   || |
Db         448 LWTSAVPNFSGLKRFGCVVYVYSQEGKLDPR----AKKGVFVGY----PNGVKGFRVWMI 499

Qy          56 --EKCSKQLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPKFSFCG 113
                 |:||         |   :: |||:  :: |  :::| :| ::| :|:    :| |   |
Db         500 EEERCS--------ISRNVVFREDV--MYKD--ILNQSTSGMSFDFPLATNRIPSFECAG 547

Qy         114 RRKGEQIYYAGPVNNPE 130
               || ::|   | |:: :
Db         548 NRKEDEISVQGGVSDDD 564


RESULT 3
Q7TUM5
ID   Q7TUM5      PRELIMINARY;      PRT;   204 AA.
AC   Q7TUM5;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Possible occludin/ELL family.
GN   PMT1867.
OS   Prochlorococcus marinus (strain MIT 9313).
OC   Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC   Prochlorococcus.
OX   NCBI_TaxID=74547;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22825698; PubMed=12917642;
RA   Rocap G., Larimer F.W., Lamerdin J., Malfatti S., Chain P.,
RA   Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,
RA   Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,
RA   Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
RA   Webb E.A., Zinser E.R., Chisholm S.W.;
RT   "Genome divergence in two Prochlorococcus ecotypes reflects oceanic
RT   niche differentiation.";
RL   Nature 424:1042-1047(2003).
DR   EMBL; BX572100; CAE22042.1; -.
KW   Complete proteome.
SQ   SEQUENCE   204 AA;  22142 MW;  6515E9FBDCB2DF0A CRC64;

  Query Match             9.6%;  Score 83;  DB 16;  Length 204;
  Best Local Similarity   24.2%;  Pred. No. 1;
  Matches   37;  Conservative   22;  Mismatches   36;  Indels   58;  Gaps    9;

Qy          13 LIFP-SCSGGGGGKAWPTHVVCSDSGLE----------VLYQSCDPLQDFGFSVEKCSK 60
               |: | : :| ||| ||   |||: : ||          |    | ||:  |  :|:
Db          36 LVTPLAVTGWGGGLAWARPVVCTTT-LEAPDANQSSKGPVEVSQCGPLESTGALIER--- 91

Qy          61 QLKSNINIRF---------GIILREDIKELFLDLALMSQ-GSSVLNFSYPICEAALPKFS 110
                       ||         |: :   : :: | :|:  | |: :: | :|
Db          92 --------RFYSWTAPYVRGVDVLHQLTDM-LGIAVAGQEGNRLMGFGFP---------- 132

Qy         111 FCGRRKGEQIYYAGPVNNPEFTIPQGEYQVLLE 143
                     : | : |       :  |  ||||||
Db         133 ------DQTIIWDG-------SAVQNTYQVLLE 152


RESULT 4
Q9LV30
ID   Q9LV30      PRELIMINARY;      PRT;  1440 AA.
AC   Q9LV30;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Emb|CAB40755.1.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Columbia;
RA   Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Columbia;
RX   MEDLINE=20363099; PubMed=10907853;
RA   Nakamura Y.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT   Sequence features of the regions of 4,251,695 bp covered by ninety P1,
RT   TAC and BAC clones.";
RL   DNA Res. 7:217-221(2000).
DR   EMBL; AB020749; BAB02023.1; -.
DR   InterPro; IPR002885; PPR.
DR   InterPro; IPR008941; TPR-like.
DR   Pfam; PF01535; PPR; 25.
DR   TIGRFAMs; TIGR00756; PPR; 24.
SQ   SEQUENCE   1440 AA;  162336 MW;  364887F23EF8B9BF CRC64;

  Query Match             9.6%;  Score 83;  DB 10;  Length 1440;
```

```
  Best Local Similarity   22.5%;  Pred. No. 11;
  Matches   27;  Conservative   22;  Mismatches   57;  Indels   14;  Gaps    3;

Qy          30 HVVCSDSGLEVLYQSCDPLQDFGFSVEKCSKQLKSNINIRFGIILREDIKELFLDLALMS 89
               | :| |  || ||   : ||| || : | |  |  :   | | |   ::|:::  :
Db         830 HALCVDGRLEELYVVVEELQDMGFKISKSSILLMLDAFARAGNIF--EVKKIYSSMKAAG 887

Qy          90 QGSSVLNFSYPI---CEAALPKFSFCGRRKGEQIYYAGPVNNPEFTIPQGEYQVLLELYT 146
                 ::  :   |   |:         |:|  :       :    | :    :  :|::||
Db         888 YLPTIRLYRMMIELLCK---------GKRVRDAEIMVSEMEEANFKVELAIWNSMLKMYT 938


RESULT 5
O58069
ID   O58069      PRELIMINARY;      PRT;   406 AA.
AC   O58069;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein PH0331.
GN   PH0331.
OS   Pyrococcus horikoshii.
OC   Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC   Pyrococcus.
OX   NCBI_TaxID=53953;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OT3;
RX   MEDLINE=98344137; PubMed=9679194;
RA   Kawarabayasi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA   Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA   Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA   Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA   Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA   Masuchi Y., Shizuya H., Kikuchi H.;
RT   "Complete sequence and gene organization of the genome of a hyper-
RT   thermophilic archaebacterium, Pyrococcus horikoshii OT3.";
RL   DNA Res. 5:55-76(1998).
DR   EMBL; AP000002; BAA29405.1; -.
DR   PIR; H71139; H71139.
DR   InterPro; IPR002881; DUF58.
DR   Pfam; PF01882; DUF58; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   406 AA;  45912 MW;  694855F337C5DD8D CRC64;

  Query Match             9.4%;  Score 81;  DB 17;  Length 406;
  Best Local Similarity   25.8%;  Pred. No. 3.8;
  Matches   39;  Conservative   27;  Mismatches   49;  Indels   36;  Gaps    9;

Qy           1 MKG--FTATLFLWTLIFPSCSGGGGGKA-WPTHVVCSDSGLEVLYQSCDPLQDFGF--SV 55
               |||  |  ||||| ::      |  |  | :|  ::     : :|:   |   ||      :
Db           4 MKGASFFVTLFLWFIVMSLIFGVPGKLAVFPLLLLI----MGLLF---DAPGDFEVEREI 56

Qy          56 EKC----SKQLKSNINIRFG-----IILREDIKELFLDLALMSQGSSVLNFSYPICEAAL 106
               ||        :::  : :| |     :::||:| : |    : | ||:|  |
Db          57 EKAQTFVGNEIEVLVRVRVGRGIGLVMVRENIPKAF----MTSSGSNVGYF--------- 103

Qy         107 PKFSFCGRRKGEQIYYAGPVNNPEFTIPQGE 137
                 |:: |||  :  |   |:    : ||: |
Db         104 --FTYPGRRSFQLSYKLLPLKRGVYEIPKTE 132


RESULT 6
Q9SIN9
ID   Q9SIN9      PRELIMINARY;      PRT;   484 AA.
AC   Q9SIN9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Putative triacylglycerol lipase.
GN   AT2G31690.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=20083487; PubMed=10617197;
RA   Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA   Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA   Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA   Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA   Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA   Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA   Salzberg S.L., Fraser C.M., Venter J.C.;
RT   "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:761-768(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Lin X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC007071; AAD24845.1; -.
DR   PIR; H84723; H84723.
DR   GO; GO:0003824; F:catalytic activity; IEA.
DR   GO; GO:0004806; F:triacylglycerol lipase activity; IEA.
DR   GO; GO:0006629; P:lipid metabolism; IEA.
DR   InterPro; IPR002921; Lipase_3.
DR   InterPro; IPR000379; Ser_estrs.
DR   Pfam; PF01764; Lipase_3; 1.
SQ   SEQUENCE   484 AA;  54923 MW;  74650CD21BB783D3 CRC64;

  Query Match             9.2%;  Score 79.5;  DB 10;  Length 484;
  Best Local Similarity   27.4%;  Pred. No. 6.8;
  Matches   34;  Conservative   18;  Mismatches   43;  Indels   29;  Gaps    5;

Qy          35 DSGLEVLYQSCDPLQDFGFSVEKCSKQLKSNINIRFGIILREDIKELFLDLALMSQGSSV 94
               :||   || | :    ||  :| | :|| | |:    :: :   :|: : ||  | |||:
Db         248 ESGFLSLYTSDESESKFG--LESCRQQLLSEIS---RLMNKYKGEEMSITLAGHSMGSSL 302

Qy          95 LN-FSYPICEAALPK----------FSFCGRRKGE----------QIYYAGPVNNPE 130
                 :| | |  | :           |||  | | |          ::      ||:|
Db         303 AQLLAYDIAELGLNRRIGKGDIPVTVFSFAGPRVGNLEFKKRCEELGVKVLRITNVNDPV 362

Qy         131 FTIP 134
                 :|
Db         363 TKLP 366


RESULT 7
Q9VSU2
ID   Q9VSU2      PRELIMINARY;      PRT;  2786 AA.
AC   Q9VSU2; Q9VSU1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   CG4821 protein.
GN   TEQUILA OR CG4821 OR CG4948 OR CG18403.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
```

```
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA   Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA   Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA   Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA   Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA   Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA   Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA   Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA   McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA   Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA   Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA   Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA   Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT   "Sequencing of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA   Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA   Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA   Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA   Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA   Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA   Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT   "Annotation of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR   EMBL; AE003553; AAF50319.3; -.
DR   HSSP; P00750; 1RTF.
DR   FlyBase; FBgn0023479; Tequila.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008061; F:chitin binding; IEA.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0005044; F:scavenger receptor activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006030; P:chitin metabolism; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0015992; P:proton transport; IEA.
DR   InterPro; IPR000194; ATPase_a/bcentre.
DR   InterPro; IPR002557; Chitin_bind_PerA.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR002172; LDL_receptor_A.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF01607; CBM_14; 15.
DR   Pfam; PF00057; ldl_recept_a; 2.
DR   Pfam; PF00530; SRCR; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00261; LDLRECEPTOR.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   SMART; SM00494; ChtBD2; 15.
DR   SMART; SM00192; LDLa; 2.
DR   SMART; SM00202; SR; 2.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
DR   PROSITE; PS01209; LDLRA_1; 1.
DR   PROSITE; PS50068; LDLRA_2; 2.
DR   PROSITE; PS00420; SRCR_1; 2.
DR   PROSITE; PS50287; SRCR_2; 2.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Protease; Serine protease.
SQ   SEQUENCE   2786 AA;  308362 MW;  2EED7A7DA5002C76 CRC64;

  Query Match             9.2%;  Score 79;  DB 5;  Length 2786;
  Best Local Similarity   26.9%;  Pred. No. 63;
  Matches   28;  Conservative   20;  Mismatches   36;  Indels   20;  Gaps    6;

Qy         12 TLIFPSCSGGGGG-KAWP----THVVCSDSGLEVLYQSCDPLQDFGFSVEKC--SKQLKS 64
              |:|  :|  |  |   :|     :| ||| | ::  |||:    |  |   |  |: | :
Db        770 TMILSACPDGTNGLHLYPYDAGKYVRCSDGG-KMSIQSCENQMAFSLSQRACRPSRLLST 828

Qy         65 NINIRFGIILREDI--------KELFLDLALMSQGSSVLNFSYP 100
                 :||    ||::        ::: :  : : :  |||  :||
Db        829 EDRVRF----REELQIQTTYSSQDIQIQQSPLKECPSVLRGNYP 868


RESULT 8
Q12139
ID   Q12139      PRELIMINARY;      PRT;  1133 AA.
AC   Q12139;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   YPR022C protein.
GN   YPR022C.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB972;
RA   Badcock K., Bowman S., Churcher C.M., Pearson D., Rajandream M.A.,
RA   Walsh S.V., Barrell B.G.;
RL   Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
```

```
RC   STRAIN=AB972;
RA   Badcock K., Churcher C.M.;
RL   Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB972;
RA   Barrell B., Rajandream M.A.;
RL   Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z71255; CAA95018.1; -.
DR   EMBL; Z49274; CAA89276.1; -.
DR   PIR; S54496; S54496.
DR   SGD; S0006226; YPR022C.
DR   InterPro; IPR007087; Znf_C2H2.
DR   Pfam; PF00096; zf-C2H2; 2.
DR   SMART; SM00355; ZnF_C2H2; 1.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 2.
KW   Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
SQ   SEQUENCE   1133 AA;  130116 MW;  77F6F3C9E3760DAA CRC64;

  Query Match             9.0%;  Score 78;  DB 3;  Length 1133;
  Best Local Similarity   25.3%;  Pred. No. 27;
  Matches   37;  Conservative   16;  Mismatches   31;  Indels   62;  Gaps    8;

Qy         46 DPLQDFGFSVEKCS-----------------KQLKS---NINIRFGII----LREDIKE 80
              |||   :  | |||                   | |||    | : ::|::       | :
Db        472 DPLTKYKISKEKCQEMFSFVPELRYVSIESIHKSLKSFWLNFHPQYGLLHKPSFHVDKQP 531

Qy         81 LFLDLALMSQGSSVLNFSY------PICEAALPKFSFCGRRKGEQIYYAGPV-----NNP 129
                |:|||:  |:| |   |      |||                  ||:     ::
Db        532 AILNLALIMTGASFLGSEYREQISDPIC-------------------GPLRWIIFSHA 570

Qy        130 EFTIPQGEYQV----LLELYTEKRST 151
              :|  |    |:    |:| | || ||
Db        571 DFQPPSKTYIIQSLLLVEGY-EKTST 595


RESULT 9
Q960G6
ID   Q960G6      PRELIMINARY;      PRT;  2409 AA.
AC   Q960G6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   SD02860p.
GN   TEQUILA OR CG4821 OR CG4948 OR CG18403.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA   Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA   Nunoo J., Pacleb J., Paragas V., Park S., Phouanenavong S., Wan K.,
RA   Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR   EMBL; AY052072; AAK93496.1; -.
DR   HSSP; P00761; 1AN1.
DR   FlyBase; FBgn0023479; Tequila.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008061; F:chitin binding; IEA.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0005044; F:scavenger receptor activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006030; P:chitin metabolism; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0015992; P:proton transport; IEA.
DR   InterPro; IPR000194; ATPase_a/bcentre.
DR   InterPro; IPR002557; Chitin_bind_PerA.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR002172; LDL_receptor_A.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR001190; Srcr_receptor.
DR   Pfam; PF01607; CBM_14; 10.
DR   Pfam; PF00057; ldl_recept_a; 2.
DR   Pfam; PF00530; SRCR; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00261; LDLRECEPTOR.
DR   PRINTS; PR00258; SPERACTRCPTR.
DR   SMART; SM00494; ChtBD2; 10.
DR   SMART; SM00192; LDLa; 2.
DR   SMART; SM00202; SR; 2.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
DR   PROSITE; PS01209; LDLRA_1; 1.
DR   PROSITE; PS50068; LDLRA_2; 2.
DR   PROSITE; PS00420; SRCR_1; 2.
DR   PROSITE; PS50287; SRCR_2; 2.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Protease; Serine protease.
SQ   SEQUENCE   2409 AA;  267618 MW;  4F014411D0B87150 CRC64;

  Query Match             9.0%;  Score 78;  DB 5;  Length 2409;
  Best Local Similarity   26.0%;  Pred. No. 68;
  Matches   27;  Conservative   21;  Mismatches   36;  Indels   20;  Gaps    6;

Qy         12 TLIFPSCSGGGGG-KAWP----THVVCSDSGLEVLYQSCDPLQDFGFSVEKC--SKQLKS 64
              |:|  :|  |  |   :|     :| ||| | ::  |||:    |  |   |  |: : :
Db        394 TMILSACPDGTNGLHLYPYDAGKYVRCSDGG-KMSIQSCENQMAFSLSQRACRPSRLVST 452

Qy         65 NINIRFGIILREDI--------KELFLDLALMSQGSSVLNFSYP 100
                 :||    ||::        ::: :  : : :  |||  :||
Db        453 EDRVRF----REELQIQTTYSSQDIHIQQSPLKECPSVLRGNYP 492


RESULT 10
Q9UY77
ID   Q9UY77      PRELIMINARY;      PRT;   394 AA.
AC   Q9UY77;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   PEPX related from CALDICELLULOSIRUPTOR SACCHAROLYTICUS.
GN   PYRAB16310 OR PAB1277.
OS   Pyrococcus abyssi.
OC   Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC   Pyrococcus.
OX   NCBI_TaxID=29292;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GE5 / Orsay;
RA   Heilig R.;
RT   "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
RT   structure and evolution.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ248288; CAB50535.1; -.
DR   PIR; A75012; A75012.
DR   InterPro; IPR002881; DUF58.
DR   Pfam; PF01882; DUF58; 1.
KW   Complete proteome.
SQ   SEQUENCE   394 AA;  44831 MW;  39C29C86B1A95245 CRC64;

  Query Match             9.0%;  Score 77.5;  DB 17;  Length 394;
  Best Local Similarity   23.4%;  Pred. No. 8.7;
  Matches   36;  Conservative   31;  Mismatches   44;  Indels   43;  Gaps    9;
```

```
Qy          1 MKG--FTATLFLWTLIFPSCSGGGGGKA-WPTHVVCSDSGLEVLYQSCDPLQDFGFSVEK 57
              |:|  | |||||  ::   |  |  | :|  ::    : ::: :      ||:||:
Db          1 MRGVSFFLTLFLWFVLLSVIFGTPGKLALFPLLIII----VGMIFDT-----PGGFTVER 51

Qy         58 --------CSKQLKSNINIRFG-----IILREDIKELFLDLALMSQGSSVLN-FSYPICE 103
                    ::::  : :|       :||:|::   |       :||:  : |:||
Db         52 EVTKRRVFVGEEVRVKLRVRVSKGIGIVILKENLPRAF-----EVKGSATFHFFTYPGKR 106

Qy        104 AALPKFSFCGRRKGEQIYYAGPVNNPEFTIPQGE 137
                ::||  ||||            |: ||: |
Db        107 EFSFEYSFVPRRKG-----------EYDIPKTE 128


RESULT 11
O23839
ID   O23839      PRELIMINARY;      PRT;   428 AA.
AC   O23839;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   S glycoprotein (Fragment).
OS   Brassica oleracea (Cauliflower).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Brassica.
OX   NCBI_TaxID=3712;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97352858; PubMed=9207151;
RA   Kusaba M., Nishio T., Satta Y., Hinata K., Ockendon D.;
RT   "Striking sequence similarity in inter- and intra-specific comparisons
RT   o f class I SLG alleles from Brassica oleracea and Brassica
RT   campestris: Implications for the evolution and recognition
RT   mechanism.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:7673-7678(1997).
DR   EMBL; D85205; BAA21939.1; -.
DR   PIR; T14529; T14529.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   InterPro; IPR001480; B_lectin.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR000858; Slocus_glycop.
DR   Pfam; PF01453; Agglutinin; 1.
DR   Pfam; PF00954; S_locus_glycop; 1.
DR   SMART; SM00108; B_lectin; 1.
DR   SMART; SM00473; PAN_AP; 1.
FT   NON_TER       1      1
FT   NON_TER     428    428
SQ   SEQUENCE   428 AA;  48724 MW;  18E04542C7293BEA CRC64;

  Query Match             9.0%;  Score 77.5;  DB 10;  Length 428;
  Best Local Similarity   25.2%;  Pred. No. 9.7;
  Matches   27;  Conservative   20;  Mismatches   49;  Indels   11;  Gaps    4;

Qy         49 QDFGFSVEKCSKQLKSNINIRFGIILREDIKELFLDLALMSQGSSVLNFSYPICEAALPK 108
              | | |  :    ::|   |:: |:          ||      |   | ::|| :    ||:
Db        149 QSFDFPTDTLLPEMKLGYNLKTGL-------NRFLTAWRNSDDPSSGDYSYKLENRELPE 201

Qy        109 FSFCGRRKGEQIYYAGPVNNPEFT-IPQGE-YQVLLELYTEKRSTVA 153
                  :  | |:: :|| |   |: ||: :    ::  :||    ||
Db        202 FYLL--KSGFQVHRSGPWNGVRFSGIPENQKLSYMVYNFTENSEEVA 246


RESULT 12
Q8ZBY9
ID   Q8ZBY9      PRELIMINARY;      PRT;   246 AA.
AC   Q8ZBY9;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative exported protein.
GN   YPO3241.
OS   Yersinia pestis
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Yersinia.
OX   NCBI_TaxID=632;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CO-92 / Biovar Orientalis;
RX   MEDLINE=21470413; PubMed=11586360;
RA   Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA   Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,
RA   Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA   Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA   Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA   Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA   Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT   "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL   Nature 413:523-527(2001).
DR   EMBL; AJ414156; CAC92476.1; -.
DR   PIR; AH0393; AH0393.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   246 AA;  27994 MW;  83FA14455C2F82BC CRC64;

  Query Match             8.9%;  Score 77;  DB 16;  Length 246;
  Best Local Similarity   29.0%;  Pred. No. 5.6;
  Matches   31;  Conservative   12;  Mismatches   30;  Indels   34;  Gaps    6;

Qy         45 CDPLQDFGFSVEK---CSKQLKSNI---NIRFGIILREDIKELFLDLALMSQGSSVLNFS 98
              | |   | |  :    ||:||  :   ::   |   | : ||:  |    ||   | |
Db         13 CLPAASFTSSASEPVAVSKELKQQLLGSSVYIQIFKEERVLELYAKL----QGEYRLVQS 68

Qy         99 YPICEAALPKFSFCG-----RRKGEQIYYAGPVNNPEFTIPQGEYQV 140
              |||||       | |     ||:|             :|  |:| |::
Db         69 YPICE-------FSGGLGPKRREG-----------DFKSPEGFYRI 96


RESULT 13
O76660
ID   O76660      PRELIMINARY;      PRT;   396 AA.
AC   O76660;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Putative GPI-anchored surface antigen SRS4.
GN   SRS4.
OS   Toxoplasma gondii.
OC   Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Sarcocystidae;
OC   Toxoplasma.
OX   NCBI_TaxID=5811;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RH;
RA   Hehl A., Boothroyd J.C.;
RT   "SRS4, a member of the SAG1 family of surface antigens in Toxoplasma
RT   gondii.";
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF078678; AAC27649.1; -.
DR   InterPro; IPR007226; SAG.
DR   Pfam; PF04092; SAG; 1.
DR   PRINTS; PR01801; SURFCEANTIGN.
SQ   SEQUENCE   396 AA;  42105 MW;  7E96A39F44D15A02 CRC64;

  Query Match             8.9%;  Score 77;  DB 5;  Length 396;
  Best Local Similarity   21.1%;  Pred. No. 10;
  Matches   37;  Conservative   22;  Mismatches   62;  Indels   54;  Gaps    6;

Qy         24 GKAWPTHVVCSDSGL--EVLYQSCDPLQDFGFSVEKCSKQLKSNINIR---------- 69
              |: |   :  ||  |  :  :  ||     |  |:  ||: | : |::
Db        170 GEEWTLQLQESDLPLTDKAFFVGCDNKAGAGKDVQTSSKECKVDFNVKARPSFVAENNVV 229

Qy         70 ----------------------------FGIILREDIKELFLDLALMSQGSSVLNFSY 99
                                          |::  :    | ||    | : : | :
Db        230 TCAYGKESNPEPLNVEMTTEMNTLTIQCGSHGVLNPKSYAATFCDL----QDTDMQNCTE 285

Qy        100 PICEAALPKF--SFCGRRKGEQIYYAGPVNNPEFTIPQGEYQVLLE-LYTEKRST 151
```

```
                 || || |  ::   |  :|   :  :  ||   || | |  |  :| | :||
Db        286 KKFEAILPSFVETWWAARSKDQ---SATLTIPEAEFPQSEQQFRLSCIYKEAQST 337


RESULT 14
Q7T2R7
ID   Q7T2R7      PRELIMINARY;      PRT;   480 AA.
AC   Q7T2R7;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Nicotinic acetylcholine receptor beta 1b subunit (Fragment).
OS   Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Takifugu.
OX   NCBI_TaxID=31033;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Jones A.K., Elgar G., Sattelle D.B.;
RT   "The nicotinic acetylcholine receptor gene family of the pufferfish,
RT   Fugu rubripes.";
RL   Genomics 0:0-0(2003).
DR   EMBL; AY299468; AAP58383.1; -.
KW   Receptor.
FT   NON_TER       1      1
SQ   SEQUENCE   480 AA;  55233 MW;  D0CD2322938E06C0 CRC64;

  Query Match             8.9%;  Score 77;  DB 13;  Length 480;
  Best Local Similarity   29.9%;  Pred. No. 13;
  Matches   23;  Conservative   14;  Mismatches   24;  Indels   16;  Gaps    4;

Qy         23 GGKAW-PTHVVCSDSGLEVLYQSCDPLQDFGFSVEKCSKQLKS------NINIRFGIILR 75
              |  | |  : ||  |::| |         | |  : || | :|       |:::: : ||
Db        115 GRVTWTPPALYCSSCGVKVEY--------FPFDWQNCSMQFRSYTYDSTEIDVQYALNLR 166

Qy         76 -EDIKELFLDLALMSQG 91
               ::|:|: || |   |
Db        167 GQEIREIQLDEAFTEGG 183


RESULT 15
Q9M9G4
ID   Q9M9G4      PRELIMINARY;      PRT;   739 AA.
AC   Q9M9G4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   F14O23.20 protein.
GN   F14O23.20.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. columbia;
RA   Liu S., Yu G., Lee J., Sakano H., Jhaveri A., Lenz C., Toriumi M.,
RA   Chin C., Chiou J., Choi E., Gonzalez A., Howng B., Koo T., Li J.,
RA   Liu A., Pham P., Vaysberg M., Altafi H., Buehler E., Chao Q., Conn L.,
RA   Conway A.B., Hansen N., Johnson-Hopson C., Khan S., Kim C., Lam B.,
RA   Nguyen M., Palm C., Shinn P., Tambunga G., Davis R., Ecker J.,
RA   Federspiel N., Theologis A.;
RT   "The sequence of BAC F14O23 from Arabidopsis thaliana chromosome 1.";
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. columbia;
RA   Theologis A.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC012654; AAF43235.1; -.
DR   PIR; G96740; G96740.
SQ   SEQUENCE   739 AA;  84156 MW;  9EDA5EEEF888BAEB CRC64;

  Query Match             8.9%;  Score 77;  DB 10;  Length 739;
  Best Local Similarity   24.0%;  Pred. No. 21;
  Matches   25;  Conservative   17;  Mismatches   30;  Indels   32;  Gaps    4;

Qy         48 LQDFGFSVEKCSKQLKSNINIRFGIILR---EDIKELFLDLALMSQGSSVLNFSYPICEA 104
              :|  |:  :|| :|::  :  ||  :|    ||:|                       |
Db        264 VQGKGYK-DKCYEQIRKAVEDRFNRLLTLVFEDLK----------------------A 298

Qy        105 ALPKFSFCGRRKGEQIYYAGPVNNPEFTIPQGEYQVLLELYTEK 148
              || :     |  |:  |  |   | : |    :|::: ||||:
Db        299 ALEEARMIGEELGDIYDYVAPCFPPRYEI----FQLMVNLYTER 338


Search completed: April 29, 2004, 15:23:13
Job time : 43 secs
```